Plank Problems, Hadamard Matrices and Lipschitz Maps

by

Oscar Adrian Ortega Moreno

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Mathematics Institute

October 2019

Contents

To my daughter Ivanna

Acknowledgments

First of all, I would like to thank my supervisor, Professor Keith M. Ball, for his invaluable guidance and constant motivation during the course of my PhD studies. Without his help this work would not have been possible.

I would also like to acknowledge the financial support I received from Mexico's National Council for Science and Technology (CONACyT) along my PhD studies at the University of Warwick.

Last but not least, I would like to thank my family and friends. Especially, I am grateful to my parents for encouraging me to always pursue my dreams. Thank you for all your love and support.

Declarations

Apart from the results given in the second chapter, obtained in collaboration with Ball and Prodromou [6], I declare that the material in this thesis is my own except where otherwise indicated or cited in the text.

Abstract

Chapter 1 proves an optimal version of the plank theorem in real Hilbert spaces. Plank problems are questions concerning coverings of convex sets by planks (regions between two parallel hyperplanes). The problem treated here is related to coverings of unit balls of real Hilbert spaces by collections of planks that are symmetric about the origin. Chapter 2 discusses a connection between two combinatorial designs: 1-factorizations and Hadamard matrices. We consider 1-factorizations of complete graphs that match a given Hadamard matrix. The existence of these factorizations is established for two well-known families of Hadamard matrices: Walsh matrices and certain Paley matrices. Chapter 3 studies Markov type properties for L_p spaces for $p \in (1, 2)$. The notion of Markov type was introduced by Ball and it describes the evolution of time-reversible Markov chains with a finite number of states on a given Metric space. Ball showed that there is striking connection between this property and the extension of Lipschitz maps. Exploiting this connection, we obtain some results concerning the extension of Lipschitz maps defined on L_p spaces with $p \in [1, 2]$.

Preface

This thesis treats various problems that lie in the fields of Discrete Geometry, Functional Analysis, Metric Geometry and Combinatorial Design Theory. Although they are not directly connected, all of them deal with the structure or distribution of finitely many points in some geometrical or combinatorial setting. The chapters are independent and can be considered as individual pieces of work. For this reason, each chapter contains its own introduction and here we just give a brief overview of the content in each of them.

In the first chapter, we study plank problems using inverse eigenvectors. A plank in a vector space is the region bounded by two parallel hyperplanes. Plank problems are concerned with coverings of closed sets by collections of planks. The classic plank problem was originally posed by Tarski in the early 1930s (in connection to the Banach-Tarski paradox). It states that if a collection of planks covers a convex body (compact convex subset of $\mathbb{R}^n$), then the sum of the widths of the planks must be at least the minimal width of the convex body they cover. Tarski proved it for the unit circle and the 3-dimensional ball. In 1951, Bang solved the problem in its full generality. At the end of his paper, Bang asked if the plank problem could be strengthened by measuring the widths of the planks with respect to the convex body that is being covered. Given a convex body K, the relative width of a plank is the width of the plank divided by the width of K in the direction normal to the plank. Bang's question can be formally stated as follows: if a convex body is covered by a collection of planks, must the sum of relative widths of the planks be at least 1?

This question is more natural than the original plank problem since it is affine invariant. Although this affine version of the plank problem remains open, Ball [2] solved it for symmetric convex bodies (Ball actually solved it for arbitrary unit balls of Banach spaces regardless of the dimension of the space). Ball's plank theorem can be seen as a generalization of the Hahn-Banach theorem, a sharp quantitative version of the uniform boundedness principle or a geometric pigeonhole principle. Ball's theorem also drew some connection between the plank problem and the coefficient problem in harmonic analysis [4]. Since then, several variants of the plank problem have emerged in complex and spherical geometry (see [10]). In this thesis, we will use inverse eigenvectors to transfer plank problems, which are purely geometrical, to the study of the extremal behavior of a certain classes of functions. Given a matrix M, w is an inverse eigenvector of M if $Mw = w^{-1}$. This method was developed by Ball [5] in his solution of the complex plank problem and later adapted by Ambrus [1] to tackle the strong polarization problem. We use inverse eigenvetors to give a new proof of a generalization of Fejes Tóth's zone conjecture [29] from 1973. This conjecture was recently solved by Jiang and Polyanskii [13] using a completely different method. We prove that for any sequence $v_1, v_2, ..., v_n$ of unit vectors in a real Hilbert space H, there exists a unit vector v in H such that

$$|\langle v_k, v\rangle| \geq \sin(\pi/2n)$$

for all k. This can be seen as a sharp version of the plank theorem for real Hilbert spaces. As a result, we obtain a unified approach to some of the most important plank problems on the real and complex setting: the classic plank problem, the complex plank problem, Fejes Tóth's zone conjecture and the strong polarization problem.

In the second chapter, we study a connection between two different combinatorial designs: 1-factorizations and Hadamard matrices. More precisely, we study 1-factorizations of complete graphs that "match" a given Hadamard matrix. With this we mean that there are restrictions for the factorizations in terms of a given Hadamard matrix. One restriction is that the edges of the k-th factor must have endpoints of opposite sign in the k-th row of the Hadamard matrix; the other, that the edges have endpoints of the same sign. We conjecture that such factorizations exist for any given Hadamard matrix. We give an integer-program formulation of the problem and show the existence of such factorizations for some well known classes of Hadamard matrices: Walsh matrices and Paley matrices of certain sizes. In the final chapter, we deal with extension of Lipschitz maps to finitely many points. We give a simple proof of a Type 2 property for L_p spaces with $p \in [1, 2]$ and we explain how this property is used to prove a strengthening of Maurey's extension theorem. We propose a possible metric analogue of this property. Recall that a stochastic matrix is a square matrix of non-negative real numbers whose rows add up to one. So we have the following conjecture.

CONJECTURE. *Let $1 < p < 2$, $x_1, \ldots, x_n \in L_p$ and A a $n \times n$ symmetric stochastic matrix. Then, for any positive integer t we have that*

$$\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^2 \leq K_p (\log n)^{\frac{2}{p}-1} t \sum_{ij} a_{ij} \|x_i - x_j\|_{L_p}^2$$

where K_p is a universal constant depending only on p.

We prove this conjecture for $p = 1$ (for an arbitrary metric space) and we get a weaker version for general p. After that, we prove that this is true when A is the transition matrix of the standard random walk on a hypercube (actually, for any transition matrix on the cube for which the transition probabilities only depend on the Hamming distance of the vertices of the cube). Finally, we give some results concerning the extension of Lipschitz maps defined on L_p spaces with $p \in [1, 2]$.

Chapter 1

An Optimal Plank Theorem

1.1 Introduction

A plank in a vector space X is the region bounded by two parallel hyperplanes. The classical plank problem, conjectured by Tarski, states that if an n-dimensional convex body is covered by a collection of planks, then the sum of the widths of the planks should be at least the minimal width of the convex body they cover. Tarski proved it for the particular cases of the unit disc and the 3-dimensional solid sphere. Bang [7] solved the problem for arbitrary convex bodies. Bang [7] also asked whether the widths of the planks could be measured with respect to the convex body that it is covered. Ball [2] answered affirmatively this affine version of the plank problem for the most interesting case: when the convex body in question is symmetric. Ball's plank theorem can be seen as a generalization of the Hahn-Banach theorem, a sharp quantitative version of the uniform boundedness principle, or a geometric pigeon-hole principle.

A plank in a normed space X is a region of the form

$$\{x \in X : |\phi(x) - m| \leq w\}$$

where ϕ is a linear functional on X^* of norm 1, m a real number, and w is a positive number. The number w is called the half-width of the plank. Ball's affine plank theorem states the following.

THEOREM 1.1.1 (The Plank Theorem [2]). *For any sequence $(\phi_k)_{k=1}^{\infty}$ of norm one functionals on a real Banach space X, $(m_k)_{k=1}^{\infty}$ a sequence of real numbers and non-negative numbers $(t_k)_{k=1}^{\infty}$ satisfying*

$$\sum_{k=1}^{\infty} t_k < 1,$$

there exists a unit vector x in X for which

$$|\phi_j(x) - m_j| > t_j$$

for every j.

The Plank theorem is obviously sharp in the sense that the unit ball of X can be covered by n non-overlapping parallel planks whose half-widths add up to 1.
In the present discussion, we are interested in the affine problem in the case that the planks covering the convex body are symmetric about the origin: so we are only interested in planks of the form

$$\{x \in X : |\phi(x)| \leq w\}$$

where ϕ is a linear functional on X^* of norm 1 and w is a positive number. In this case, Ball's plank theorem states the following.
For any sequence $(\phi_k)_{k=1}^{\infty}$ of norm one functionals on a (real) Banach space X and non-negative numbers $(t_k)_{k=1}^{\infty}$ satisfying

$$\sum_{k=1}^{\infty} t_k < 1,$$

there exists a unit vector x in X for which

$$|\phi_j(x)| > t_j$$

for every j.

For an arbitrary Banach space, the condition that the sequence of positive numbers $(t_k)_{k=1}^{\infty}$ add up to at most 1 is sharp. This can be seen by taking the space X to be ℓ_1 and the collection ϕ_i to be the standard basis vectors in ℓ_∞. For other spaces, such as Hilbert spaces, one might expect to be able to improve upon this condition. Ball [5] proved that for *complex* Hilbert spaces it is possible to beat any sequence for which $\sum_k t_k^2 = 1$.

THEOREM 1.1.2 (Complex Plank Theorem [5]). *For any sequence $v_1, v_2, \dots v_n$ of unit vectors in a complex Hilbert space and positive real numbers $t_1, t_2, \dots, t_n$ satisfying*

$$\sum_{k=1}^{n} t_k^2 = 1$$

there exists a unit vector $z \in \mathbb{R}^n$ such that

$$|\langle v_k, z\rangle| \geq t_k$$

for all k.

On the other hand, for *real* Hilbert spaces, this is clearly not possible. Consider $2n$ vectors $v_1, v_2, \dots, v_{2n}$ in $\mathbb{R}^2$ equally spaced around the circle: (n vectors and their negatives). For any unit vector v in $\mathbb{R}^2$ there is a i such that

$$|\langle v_i, v\rangle| \leq \sin(\pi/2n).$$

This simple statement is connected to a conjecture by Fejes Tóth which was positively answered, about two years ago, by Jiang and Polyanskii in [13]. A zone of width w is the set of points in the unit sphere at spherical distance $w/2$ of a given great circle. In 1973, Fejes Tóth conjectured that if a collection of zones of equal width covers the unit sphere then the angular width of the zones should be at least π/n.

A zone of spherical width w associated to the great circle $\mathbb{S}^2 \cap v^\top$, for a given unit vector v (see Figure 1.1), is the set given by

$$\{x \in \mathbb{S}^2 : |\langle v, x\rangle| \leq \sin(w/2)\}.$$

With this notation, Fejes Tóth conjecture can be restated and generalized as an optimal plank theorem for real Hilbert spaces.

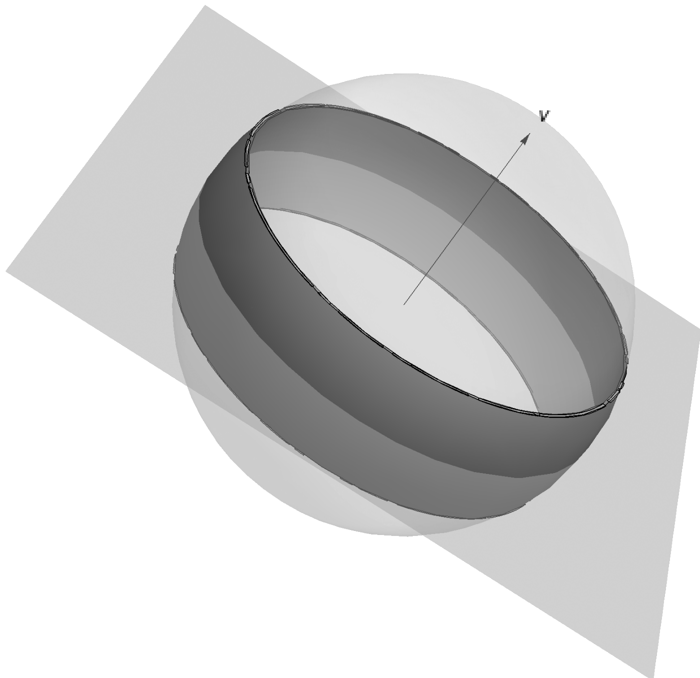


Figure 1.1: A zone of angular width $\pi/10$. The vector v is represented by a black arrow in the picture and the great circle is the intersection of the hyperplane $v^\top$ with the sphere. The zone is the set of points at angular distance $\pi/5$ of the great circle.

THEOREM 1.1.3. *For any sequence $v_1, v_2, \ldots, v_n$ of unit vectors in a real Hilbert space H, there exists a unit vector v in H such that*

$$|\langle v_k, v\rangle| \geq \sin(\pi/2n)$$

for all k.

Jiang and Polyanskii [13] solved this conjecture for arbitrary collection of zones (not necessarily having all the same width). They used the classic machinery to solve plank problems: a discrete optimization as in the classic proof of Bang's lemma [7] followed by an additional innovative inductive argument. The purpose of this chapter is to give a new proof of Fejes Tóth's zone conjecture [29] using a completely different method: Inverse Eigenvectors. We also give a different poof of the classical plank problem. As a result, we obtain a unified approach to some of the most important plank problems on the real and complex setting: the classic plank problem, the complex plank problem, Fejes Tóth's zone conjecture and the strong polarization problem.

Inverse eigenvectors allow us to transfer a purely geometrical problem to the study of the extremal behavior of certain class of functions. In the case of the complex plank problem, these are complex polynomials. In the case of Fejes Tóth conjecture and the classic plank problem, these are trigonometric polynomials.

The basic strategy in the proof of Theorem 1.1.3 is the strategy followed by Ball in the proof the complex plank problem, but there is a fundamental difference. The main ingredient of the proof of Theorem 1.1.2 has no analogue in the real case. In [5], Ball studies the behaviour of a complex polynomial locally around 1 and, with the aid of the maximum modulus principle, manages to jump away from 1 to a point in the unit disk where this polynomial has large absolute value. In contrast, the proof of Theorem 1.1.3 relies on the extremal properties of trigonometric polynomials to produce this jump.

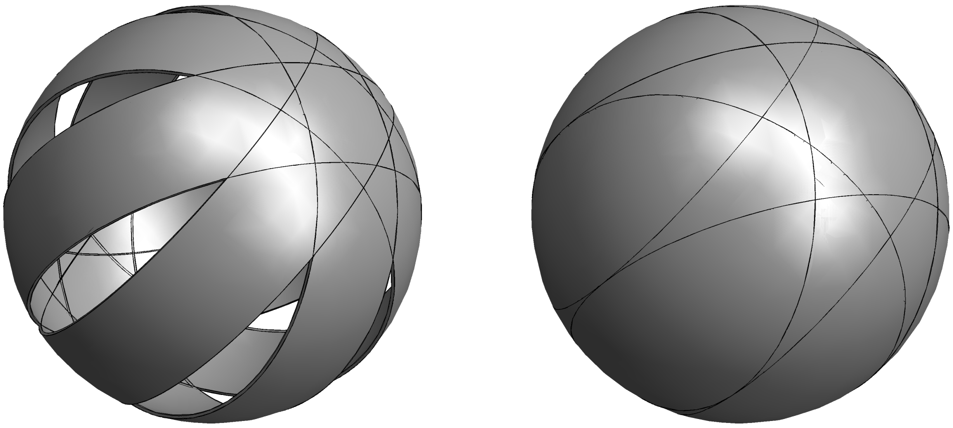


Figure 1.2: The extremal cases of Theorem 1.1.3 are sets of n vectors equally spaced around a circle $V \cap S_H$ for some 2-dimensional subspace V of H. Given such a set, the left-hand side of the figure shows sub-optimal zones, obtained from this sets, with half-width $1/n$ corresponding to the classic plank problem. On the right-hand side, the figure shows the zones with the optimal width $\sin(\pi/2n)$ corresponding to Theorem 1.1.3.

For the rest of the discussion, we shall work with the following rescaled version of Theorem 1.1.3 which will suit our purposes better. We also assume that $n \geq 2$ so as to eliminate from the discussion the trivial case when $n = 1$.

Theorem $1.0.3'$. For any sequence $v_1, v_2, \ldots, v_n$ of unit vectors in a real Hilbert space H, there exists a vector $v \in H$ of norm $\sqrt{n}$ for which

$$|\langle v_k, v \rangle| \geq \sqrt{n} \sin(\pi/2n)$$

for all k.

1.2 Inverse eigenvectors

In this section we introduce the notion of inverse eigenvectors. An inverse eigenvector of a matrix M is a vector x satisfying the equation $Mx = x^{-1}$ where x^{-1} is the inverse of x componentwise.

Inverse eigenvectors arose naturally in the solution of the complex plank problem.

In his paper [5], Ball transforms the complex plank problem to a problem concerning the location of inverse eigenvectors of a complex Gram matrix. Seven years later, Leung, Li and Rakesh [18] reformulated the problem of finding the polarization constant of $\mathbb{R}^n$ in terms of inverse eigenvectors and described the structure of the inverse eigenvectors for real positive symmetric matrices.

The term *inverse eigenvector* for a vector x satisfying $Mx = x^{-1}$ turns up for the first time in [1], where Ambrus used the methods in [5] to reformulate the strong polarization problem as a geometric question concerning the location of inverse eigenvectors and managed to solve the problem for the planar case. The treatment and adaptations of [5] and the definition of inverse eigenvectors that are presented here are due to Ambrus in [1] on his work on the strong polarization problem (see also [18]). In order to motivate the definition of inverse eigenvector, let us go back to our question.

Our problem consists of finding a vector v of norm $\sqrt{n}$ which has large inner product with all the vectors $v_1, v_2, \ldots, v_n$. An obvious candidate for this vector v would be one for which $\min_k |\langle v_k, v\rangle|$ is maximal. However, there seems to be no simple way to either manipulate or obtain useful information from this maximal condition. Instead we choose a unit vector v for which the product $\prod_i |\langle v_i, v\rangle|$ is maximal, hoping that each of the factors will be large enough to get the desired inequality. For the product, we can use simple analytic tools to study the points for which it is locally extremal. Luckily, the structure of these local optimisers can be described concisely as the following proposition shows. The following proposition can be found as discussion in the last paragraph of page 2863 in [18] or as proposition 1.16 in page in [1].

PROPOSITION 1.2.1 ([18],[1]). *Let $v_1, v_2, \ldots, v_n$ be a sequence for unit vectors in a real Hilbert space H. Suppose that v is vector of norm $\sqrt{n}$ chosen so as to maximize*

$$\prod_{k=1}^{n} |\langle v_k, v\rangle|.$$

Then,

$$v = \sum_{k=1}^{n} \frac{1}{\langle v_k, v \rangle} v_k \tag{1.2.1}$$

Proof. Since v is a stationary point, by the method of Lagrange multipliers, the gradients of the objective function and the constraint should be scalar multiples of one another. Hence, there exists a real number λ such that

$$v = \lambda \sum_{k=1}^{n} \frac{v_k}{\langle v_k, v \rangle} \prod_{k=1}^{n} | \langle v_k, v \rangle |. \tag{1.2.2}$$

This gives equation (1.2.1) up to a constant and taking inner product with v shows that the constant must be 1. $\square$

Denote by H the associated to the sequence of unit vectors $(v_k)_{k=1}^n$, that is, $H_{ij} = \langle v_i, v_j \rangle$, and let w be the vector in $\mathbb{R}^n$ given by

$$w_k = \frac{1}{\langle v_k, v \rangle} \tag{1.2.3}$$

for all k. Then w satisfies

$$(Hw)_j = \sum_{i=1}^{n} h_{ji} w_i = \langle v_j, \sum_{i=1}^{n} w_i v_i \rangle = \langle v_j, v \rangle = \frac{1}{w_j}.$$

Therefore, w satisfies the following equation $Hw = w^{-1}$ where w^{-1} is defined as the inverse of the vector w componentwise, i.e.

$$w^{-1} = \left(\frac{1}{w_1}, \ldots, \frac{1}{w_n} \right).$$

This observation leads us naturally to the following definition.

DEFINITION 1.2.2 ([1]). *Let M be a $n \times n$ matrix. We say that w is an inverse eigenvector of M if*

$$Mw = w^{-1} \tag{1.2.4}$$

In their paper [18], Leung, Li and Rakesh describe the structure of inverse eigenvectors for real Gram matrices. This is summarized in the following proposition.

PROPOSITION 1.2.3 ([18]). *Let H be a $n \times n$ real Gram matrix then:*

a) there is at most one inverse eigenvector in each quadrant of $\mathbb{R}^n$,

b) there is an inverse eigenvector in a quadrant Q of $\mathbb{R}^n$ if and only if

$$Q \cap \mathit{Ker}(H) = \{0\},$$

c) there is an inverse eigenvector in Q if and only if $\prod_i |w_i|$ has a maximum in Q. Moreover, the maximizer in Q is the unique inverse eigenvector.

Proposition 1.2.3 shows that there are at most 2^n inverse eigenvectors for a given real Gram matrix H, in contrast with the complex case, where the equation $H\bar{z} = z^{-1}$ has infinitely many solutions.
On the other hand, given an inverse eigenvector w of H, one can set

$$v = \sum_{k=1}^{n} w_k v_k.$$

It is clear that v would satisfy equations (1.2.1) and (1.2.3). Theorem $1.0.3'$ is thus a consequence of the following.

THEOREM 1.2.4. *Let H be a real Gram matrix of unit vectors. Then, there exists an inverse eigenvector w of H for which*

$$\|w\|_{\ell^\infty} \leq n^{-\frac{1}{2}} \csc(\pi/2n).$$

If one were to prove the classic plank theorem for Hilbert spaces using inverse eigenvectors, one would only have to show the following weaker version of Theorem 1.2.4.

THEOREM 1.2.5. *Let H be a* real *Gram matrix of unit vectors. Then, there exists an inverse eigenvector w of H for which*

$$\|w\|_{\ell^\infty} \leq \sqrt{n}.$$

This is an analogue of Bang's lemma in [7] (also see [2] for a proof of Bang's lemma in the form described here). To see this, we rewrite Theorem 1.2.5 as follows.

Theorem . Let H be a *real* Gram matrix of unit vectors . Then, there exists a vector w such that $|w_i| \leq \sqrt{n}$ for all i and

$$w_j \sum_k H_{jk} w_k = 1$$

and recall that Bang's lemma states slightly more than the following.

Theorem . Let H be a *real* Gram matrix of unit vectors. Then, there exists a vector of signs $\varepsilon \in \{-1, 1\}^n$ such that

$$\varepsilon_j \sum_k H_{jk} \varepsilon_k \geq \frac{1}{n}$$

We will give a simple proof of Theorem 1.2.5 using inverse eigenvectors. Actually, we will prove something stronger.

THEOREM 1.2.6. *Let H be a* real *$n \times n$ Gram matrix of unit vectors. Then, there exists an inverse eigenvector w of H for which*

$$\|w\|_H = \sup_{\|x\|=1} (wx)^\top H(wx) \leq \sqrt{n-1}$$

where wx is just the coordinate-wise product of the vectors w and x.

In terms of a sequence of vectors in a Hilbert space this says that if $v_1, \ldots, v_n$ is a sequence of unit vectors in a real Hilbert space H, then there exists a vector w in

$\mathbb{R}^n$ such that

$$\left\|\sum x_k w_k v_k\right\|_H \leq \sqrt{n-1}\,\|x\|_2$$

for all vectors x in $\mathbb{R}^n$ and

$$\left|\langle v_k, \sum w_k v_k\rangle\right| = \frac{1}{|w_k|} \geq \frac{1}{\sqrt{n-1}}.$$

This resembles a plank-type theorem of Nazarov [26], (also stated in [5]) that states the following.

Theorem . Let f_i be unit functions in L_1 which satisfy

$$\left\|\sum a_j f_j\right\| \leq M\,\|a\|_{\ell^2}$$

for some M and all $a \in \ell^2$. Let t_j be a sequence of positive numbers with $\sum_j t_j^2 = 1$. Then there is a function $g \in L_\infty$ with norm at most $15M^2$ and

$$|\langle f_j, g\rangle| \geq t_j$$

for every j.

1.3 The Proof of Theorem 1.2.4

To find a suitable eigenvector w notice that w defined as in equation (1.2.3) is a local extremal point for the function

$$\prod_{k=1}^{n} |\langle v_k, v\rangle|,$$

which in terms of w is given by

$$\prod_{k=1}^{n} \frac{1}{|w_k|}$$

subject to the constraint

$$\|v\|^2 = \left\|\sum w_k v_k\right\|^2 = w^\top H w = n$$

In the light of Proposition 1.2.1, this would suggest that we try to find a vector w to minimize $\prod |w_k|$ subject to the constraint $w^\top H w = n$. Unfortunately, this minimum is always 0. To deal with this problem, we choose u so as to maximize $\prod |u_k|$ subject to the constraint $u^\top H^{-1} u = n$, in the hope that the maximum would be converted into a minimum of the original problem via the natural bijection between the inverse eigenvectors of H and H^{-1}. That is, if u is an inverse eigenvector of H^{-1}, then $w = u^{-1}$ is an inverse eigenvector of H. We will use the following lemma in [1] which is a slight variation of Lemma 7 in [5].

LEMMA 1.3.1 ([1]). *Suppose that H is a real Gram matrix of unit vectors and w is a vector for which*

$$\prod_{k=1}^{n} |w_k|$$

is locally extremal subject to the condition

$$w^\top H w = n. \tag{1.3.1}$$

Then, w is an inverse eigenvector for H.

Lemma 1.3.1 yields a vector u for which $\prod |u_k|$ is *maximal* subject to the constraint $u^\top H^{-1} u = n$. Set $w = u^{-1}$. Thus, w is an inverse eigenvector of H. Moreover, since u has being selected as to maximize $\prod |u_k|$ we have that if c is a vector such that $\prod |c_k| = 1$, then

$$\prod |c_k u_k| = \prod |u_k|$$

and therefore

$$\sum_{jk} c_j u_j H^{-1}_{jk} c_k u_k \tag{1.3.2}$$

The problem is to show $\|w\|_\infty \leq n^{-\frac{1}{2}} \csc(\pi/2n)$.

1.4 The final transformation

In this section, we make a final transformation of the statement of Theorem 1.1.3 and then prove it. Define a matrix M by

$$m_{jk} = w_j H_{jk} w_k$$

for all j, k. M is a positive matrix and its inverse is given by

$$M_{jk}^{-1} = u_j H_{jk}^{-1} u_k.$$

Observe that

$$m_{kk} = |w_k|^2.$$

If we denote by $\mathbf{1}$ the vector whose entries are all equal to 1, then

$$\begin{aligned}
(M\mathbf{1})_j &= \sum_k w_j H_{jk} w_k \\
&= w_j \sum_k H_{jk} w_k \\
&= w_j (Hw)_j = 1
\end{aligned}$$

where the last identity is guaranteed by the fact that w is an inverse eigenvector of H. Finally, the optimal condition 1.3.2 can be restated in terms of the matrix M using the following identity

$$c^\top M^{-1} c = \sum_{jk} c_j u_j H_{jk}^{-1} c_k u_k.$$

Hence, to prove Theorem 1.2.4 it suffices to show the following.

LEMMA 1.4.1. *Suppose that M is a symmetric positive matrix satisfying*

- $M\mathbf{1} = \mathbf{1}$, *and*
- *whenever c is a vector such that $\prod |c_k| = 1$, then*

$$c^\top M^{-1} c \geq n.$$

Then $m_{kk} \leq n^{-1} \csc^2(\pi/2n)$ for all k.

In the same way Theorem 1.2.6 reduces to the following.

LEMMA 1.4.2. *Suppose that $n \geq 2$ and M is an $n \times n$ symmetric positive matrix satisfying*

- $M\mathbf{1} = \mathbf{1}$, *and*
- *whenever c is a vector such that $\prod |c_k| = 1$, then*

$$c^\top M^{-1} c \geq n.$$

Then $\|M\|_2 \leq n - 1$.

We will first give the proof for Lemma 1.4.2 and make some useful remarks that lead us to the proof of Lemma 1.4.1.

Remark 1.4.3. It must be pointed out that the proof of Lemma 1.4.2 follows the same lines as the proof of Ambrus [1] of the strong polarization problem in the planar case: the contribution here is more a refinement of the proof by using derivatives and Bernstein's inequality which potentially could be applied to a variety of classes of functions that satisfy Bernstein-type inequalities.

Proof of Lemma 1.4.2. First notice that if we let $c = Mb$ then the second condition of the lemma can be restated as follows: $\prod |(Mb)_k| = 1$ implies

$$b^\top Mb \geq n.$$

Or equivalently, for any b with

$$b^\top Mb = n,$$

$\prod |(Mb)_k| \leq 1$. The proof consists of looking at 2-dimensional slices of the ellipsoid defined by

$$\mathcal{E} = \{x : x^\top Mx = n\}.$$

So we will "cut" $\mathcal{E}$ with subspaces of dimension 2 of $\mathbb{R}^n$ which contain the vector $\mathbf{1}$. Thus, given a vector $v \in \mathcal{E}$ orthogonal to $\mathbf{1}$, we let H_v be the 2 dimensional subspace spanned by $\mathbf{1}$ and v,

$$H_v = \text{span}\{\mathbf{1}, v\}. \tag{1.4.1}$$

We denote by $\mathcal{E}_v$ the ellipse we get by intersecting $\mathcal{E}$ and H_v,

$$\mathcal{E}_v = \mathcal{E} \cap H_v. \tag{1.4.2}$$

Notice that we can parameterize the ellipse $\mathcal{E}_v$ as follows: given an angle $\theta \in [0, 2\pi]$ we define

$$v_\theta = \cos\theta\, \mathbf{1} + \sin\theta\, v. \tag{1.4.3}$$

Any vector in $\mathcal{E}_v$ is of the form (1.4.3) for some $\theta \in [0, 2\pi]$ and every vector $v_\theta \in \mathcal{E}_v$ for every $\theta \in [0, 2\pi]$. Hence,

$$\mathcal{E}_v = \{v_\theta : \theta \in [0, 2\pi]\}.$$

Define the trigonometric polynomial T_v by

$$\begin{aligned} T_v(\theta) &= \prod_{j=1}^{n} (Mv_\theta)_j \\ &= \prod_{j=1}^{n} (\cos\theta + (Mv)_j \sin\theta) \end{aligned}$$

Notice that $T_v(0) = 1$. We now compute the first and second derivatives of T_v at 0. For any θ such that $T_v(\theta)$ is not 0 we have

$$\frac{T_v'(\theta)}{T_v(\theta)} = -\sum_{j=1}^{n} \frac{\sin\theta - (Mv)_j \cos\theta}{\cos\theta + (Mv)_j \sin\theta} \tag{1.4.4}$$

Evaluating equation (1.4.4) at 0 and recalling that M is symmetric, $M\mathbf{1} = \mathbf{1}$ and v is orthogonal to $\mathbf{1}$, we see that

$$T_v'(0) = \sum_{j=1}^{n} (Mv)_j = \mathbf{1}^\top M v = \mathbf{1}^\top v = 0. \tag{1.4.5}$$

Taking derivatives on both sides of equation (1.4.4) yields

$$\frac{T_v''(\theta) T_v(\theta) - (T_v'(\theta))^2}{T_v(\theta)^2} = -\sum_{j=1}^{n} \frac{1 + (Mv)_j^2}{(\cos\theta + (Mv)_j \sin\theta)^2}. \tag{1.4.6}$$

Thus replacing $T_v(0) = 1$ and $T_v'(0) = 0$ in equation (1.4.6), we get

$$|T_v''(0)| = n + \|Mv\|^2 .$$

We are now in a position to apply the following well known inequality for trigonometric polynomials.

Theorem (Bernstein's Inequality). Let $\mathcal{T}_n$ be the set of trigonometric polynomials of degree at most n. If $T \in \mathcal{T}_n$, then

$$\|T'\|_\infty \leq n \, \|T\|_\infty \tag{1.4.7}$$

where $\|T\|_\infty$ denotes the uniform norm of T on $[0, 2\pi]$.

For a proof of Bernstein's Inequality, we refer the reader to [11] page 178. Applying Bernstein's inequality twice, we get the following inequality for the second derivative of T_v,

$$\|T_v''\|_\infty \leq n^2 \, \|T_v\|_\infty \,. \tag{1.4.8}$$

Since $v_\theta \in \mathcal{E}$,

$$|T_v(\theta)| = \prod_{j=1}^{n} |(Mv_\theta)_j| \leq 1$$

for all θ and thus

$$\|T_v\|_\infty \leq 1$$

for all $v \in \mathcal{E}$. Hence by inequality (1.4.8),

$$n + \|Mv\|^2 = |T_v''(0)| \leq \|T_v''\|_\infty \leq n^2$$

for all $v \in \mathcal{E}$ orthogonal to $\mathbf{1}$. Therefore,

$$\|Mv\|^2 \leq n(n-1) \tag{1.4.9}$$

for all $v \in \mathcal{E}$ orthogonal to $\mathbf{1}$.

Let $v \in \mathcal{E}$ be an eigenvector orthogonal to $\mathbf{1}$ associated to the possible largest eigenvalue λ. For this eigenvector v we have that

$$\|Mv\|^2 = v^\top M^\top Mv = \lambda v^\top Mv = \lambda n$$

and hence by (1.4.9),

$$\lambda \leq n - 1.$$

The norm $\|M\|_2$ is the maximum of 1 and λ which, in either case, is less than or equal to $n - 1$. $\square$

Remark 1.4.4. To get the classic result for Hilbert spaces, observe that

$$m_{kk} = e_k^\top M e_k \leq \|M\|_2 \leq n - 1 < n.$$

where e_k is the k-th canonical vector. This would give a proof for Lemma 1.2.5. However, one can try to make a better selection of the vector v so as to get a much better estimate of m_{kk} for all k. In other words, we could select the 2 dimensional slice of $\mathcal{E}$ more carefully so that we get a better bound for m_{kk} for all k. The natural choice of 2 dimensional subspace to cut $\mathcal{E}$ so as to get a better estimate for m_{kk} would be

$$H = \{x\mathbf{1} + ye_k : x, y \in \mathbb{R}\}.$$

However e_k is not orthogonal to $\mathbf{1}$ so we project it into the space orthogonal to $\mathbf{1}$. Doing so and normalizing so that the projection belongs to the ellipsoid $\mathcal{E}$, we get that H is equal to $H_{v_k} = \text{span}\{\mathbf{1}, v_k\}$ where

$$v_k = \frac{ne_k - \mathbf{1}}{\sqrt{nm_{kk} - 1}} \tag{1.4.10}$$

Then,

$$\|Mv_k\|^2 = \frac{n^2 \|Me_k\|^2 - n}{nm_{kk} - 1} \tag{1.4.11}$$

and

$$\|Me_k\|^2 = \sum_{j=1}^{n} m_{kj}^2 = m_{kk}^2 + \sum_{j \neq k} m_{kj}^2. \tag{1.4.12}$$

By inequality (1.4.9)

$$\|Mv_k\|^2 \leq n(n-1). \tag{1.4.13}$$

Using equations (1.4.11) and (1.4.12) in (1.4.13) and rearranging, we obtain

$$n^2 m_{kk}^2 + n^2 \sum_{j \neq k} m_{kj}^2 - n \leq n(n-1)(n m_{kk} - 1).$$

On the other hand, we know that $\sum_{j \neq k} m_{kj} = 1 - m_{kk}$ since $M\mathbf{1} = \mathbf{1}$ and so

$$\frac{(m_{kk}-1)^2}{n-1} \leq \sum_{j \neq k} m_{kj}^2$$

Hence,

$$n^2 m_{kk}^2 + n^2 \frac{(m_{kk}-1)^2}{n-1} - n \leq n(n-1)(n m_{kk} - 1).$$

Simplifying the above inequality yields

$$n^2 m_{kk}^2 - 2n m_{kk} + 1 \leq (n-1)^2 (n m_{kk} - 1).$$

Substituting $t = n m_{kk}$ we get

$$t^2 - 2t + 1 \leq (n-1)^2 (t-1)$$

which is true if and only if

$$1 \leq t \leq 1 + (n-1)^2$$

which corresponds to

$$\frac{1}{n} \leq m_{kk} \leq \frac{1 + (n-1)^2}{n} \tag{1.4.14}$$

For all k.

The leftmost inequality of (1.4.14) is the minimal condition m_{kk} should satisfy since M is positive and $M\mathbf{1} = \mathbf{1}$. The right hand side gives an improvement over the classic plank theorem for Hilbert spaces. In other words, this grantees that if $v_1, \ldots, v_n$ is a set of unit vectors on a Hilbert space H there exists a unit vector

$v \in H$ such that

$$|\langle v_i, v \rangle| \geq \frac{1}{\sqrt{1 + (n-1)^2}}.$$

However, this is far form being optimal. In fact, this is asymptotically equivalent to the classic result. We will give a slightly different argument for the optimal bound.

Proof of Lemma 1.4.1. Notice that if we let $c = Mb$ then the second condition of lemma 1.4.1 states that if $\prod |(Mb)_j| = 1$,

$$b^\top Mb \geq n.$$

Let us assume, for a contradiction, that one of the diagonal entries is too large. Thus, assume that there exists k such that

$$m_{kk} > \frac{1}{n \sin^2(\pi/2n)}. \tag{1.4.15}$$

We will show that there is a vector b such that $\prod |(Mb)_j| \geq 1$, but

$$b^\top Mb < n.$$

Consider the following vector

$$v_k^{(\alpha)} = -\sqrt{\alpha} v_k$$

where v_k is defined as in (1.4.10) and

$$\alpha = \frac{\cot^2(\pi/2n)}{nm_{kk} - 1}$$

The first thing we should notice is that $\alpha \in (0, 1)$. In fact, from (1.4.15) it immediately follows that

$$nm_{kk} - 1 > \frac{1}{\sin^2(\pi/2n)} - 1 = \cot^2(\pi/2n).$$

For each $\theta \in [0, 2\pi]$ define

$$v_\theta^{(\alpha)} = \cos\theta \mathbf{1} + \sin\theta\, v_k^{(\alpha)}. \tag{1.4.16}$$

It is easy to see that $v_\theta^{(\alpha)}$ is just a parametrisation of a $2-$dimensional ellipsoid *inside* $\mathcal{E}$. In fact,

$$v_\theta^{(\alpha)\top} M v_\theta^{(\alpha)} = n(\cos^2\theta + \alpha\sin^2\theta) \tag{1.4.17}$$

since $\mathbf{1}^\top M = \mathbf{1}$, $v_k^{(\alpha)}$ is orthogonal to $\mathbf{1}$ and $v_k^{(\alpha)\top} M v_k^{(\alpha)} = \alpha n$. Thus, if $\theta \in [0, 2\pi) \setminus \{0, \pi\}$, then

$$v_\theta^{(\alpha)\top} M v_\theta^{(\alpha)} < n,$$

and

$$v_\theta^{(\alpha)\top} M v_\theta^{(\alpha)} = n$$

if and only if $\theta = 0$ or $\theta = 0$. Define the trigonometric polynomial $T_{v_k^{(\alpha)}}$ by

$$T_{v_k^{(\alpha)}}(\theta) = \prod_{j=1}^{n} (M v_\theta^{(\alpha)})_j$$

or equivalently

$$T_{v_k^{(\alpha)}}(\theta) = \prod_{j=1}^{n} \left(\cos\theta + (M v_k^{(\alpha)})_j \sin\theta\right). \tag{1.4.18}$$

Notice that

$$(M v_k^{(\alpha)})_k = -\sqrt{\alpha}(M v_k)_k = -\sqrt{\alpha}\frac{e_k^\top(nMe_k - \mathbf{1})}{\sqrt{nm_{kk} - 1}} = -\sqrt{\alpha}(nm_{kk}-1) = \cot\left(\frac{\pi}{2n}\right)$$

and so the k-th factor of $T_{v_k^{(\alpha)}}$ is equal to 0 if and only if

$$\cos\theta = \cot\left(\frac{\pi}{2n}\right)\sin\theta$$

which happens if and only if $\theta = \frac{\pi}{2n}$ or $\pi + \frac{\pi}{2n}$. Hence, $T_{v_k^{(\alpha)}}$ has a root at $\theta = \frac{\pi}{2n}$ and $\pi + \frac{\pi}{2n}$. Expanding the product we get

$$T_{v_k^{(\alpha)}}(\theta) = \cos^n\theta + \sum_j (Mv_k^{(\alpha)})_j \cos^{n-1}\theta\sin\theta + \sin^2\theta\;\psi(\theta)$$

where ψ is a trigonometric polynomial of degree at most $n-2$. On the other hand,

$$\sum_j (Mv_k^{(\alpha)})_j = \mathbf{1}^\top M v_k^{(\alpha)} = \mathbf{1}^\top v_k^{(\alpha)} = 0$$

since $\mathbf{1}^\top M = \mathbf{1}$, $v_k^{(\alpha)}$ is orthogonal to $\mathbf{1}$. Therefore,

$$T_{v_k^{(\alpha)}}(\theta) = \cos^n\theta + \sin^2\theta\;\psi(\theta). \tag{1.4.19}$$

It is easy to see that $\cos n\theta$ is of the form (1.4.19); thus, taking the difference of $T_{v_k^{(\alpha)}}(\theta)$ and $\cos n\theta$ we get

$$\begin{aligned} Q(\theta) &:= T_{v_k^{(\alpha)}}(\theta) - \cos n\theta \\ &= \sin^2\theta\;\psi(\theta) \end{aligned}$$

where ψ is a trigonometric polynomial of degree at most $n-2$.

Observe that Q has roots at 0 and π, where $T_{v_k^{(\alpha)}}$ and $\cos n\theta$ are both 1, and at $\frac{\pi}{2n}$ and $\pi + \frac{\pi}{2n}$, where both functions are equal to 0.

For a contradiction, let us assume that

$$|T_{v_k^{(\alpha)}}(\theta)| < 1 \tag{1.4.20}$$

for all $\theta \in I_n = [\frac{\pi}{n}, \frac{(n-1)\pi}{n}] \cup [\frac{(n+1)\pi}{n}, \frac{(2n-1)\pi}{n}]$. The extrema of $\cos n\theta$ on $[0, 2\pi)$ are located at $\theta_k = \frac{k\pi}{n}$ for $k \in \{0, \dots, 2n-1\}$ so

$$\text{sgn } Q(\theta_k) = (-1)^{k+1}.$$

Thus, by the intermediate value theorem, for each $k \in \{1, \dots, n-2\} \cup \{n+1, \dots, 2n-2\}$ there is a $\varphi_k \in (\frac{k\pi}{n}, \frac{(k+1)\pi}{n})$ such that

$$Q(\varphi_k) = 0.$$

This gives us $2n-4$ additional roots of Q in I_n. Hence, Q has $2n$ distinct roots on the interval $[0, 2\pi)$: 0, π, $\frac{\pi}{2n}$, $\pi + \frac{\pi}{2n}$ and the $2n-4$ roots in I_n. However, Q is the product of $\sin^2(\theta)$ and a trigonometric polynomial of degree at most $n-2$ so it could not have more than $2n-2$ distinct roots. Therefore, there exists $\theta \in I_n$ such that

$$|T_{v_k^{(\alpha)}}(\theta)| = \prod_j |(Mv_\theta^{(\alpha)})_j| \geq 1$$

and ${v_\theta^{(\alpha)}}^\top M v_\theta^{(\alpha)} < n$ which is a contradiction to the second condition of Lemma 1.4.1.

$\square$

Chapter 2

Hadamard Matrices and 1-Factorization

2.1 Introduction

In this chapter we study a connection between two types of combinatorial designs: 1-factorizations and Hadamard matrices. There is not a formal definition of combinatorial design theory. However, it can be thought as the study of arrangements of finite sets into sub-classes so that they satisfy a *balance* configuration or *geometrical* condition. The central objective of combinatorial design theory is to establish whether designs of certain kinds exist. Even though it emerged as recreational mathematics, the study of designs as a mathematical discipline flourished due to their applications in the design of statistical experiments, tournament scheduling, lotteries, mathematical biology, algorithm design and analysis, networking, group testing, and cryptography. Hadamard matrices are one of the most important combinatorial design. They have a variety of applications: some of them are error-correcting codes and balanced repeated replication in statistics. A Hadamard matrix of order n is a $n \times n$ orthogonal matrix H with entries ± 1. It is easy to show that the order of a Hadamard matrix must be a multiple of 4 using the orthogonality condition. It is a famous open conjecture, due to J. Hadamard in 1893, that this is a sufficient condition for a Hadamard matrix to exist. In following paragraphs, we discuss 1-factorizations of complete graphs that match a given Hadamard matrix. We solve this conjecture for some well-known families of Hadamard matrices. First we start with some basic definitions and notation.

A *graph* is a pair $G = (V, E)$, where V is a set, called the *vertex set* of G, and E is a subset of $\binom{V}{2} = \{\{x, y\} : x, y \in V\}$ of unordered pairs of G, called the *edge set* of G. The elements of V and E are called vertices and edges, respectively. The elements of an edge e are called the *endpoints* of e. Two different edges of G are said to be *independent* if they do not share an endpoint. A set of edges $F \subset E$ spans G if the set of endpoints of the edges in F includes all of the vertices of G. A *complete graph* is a graph in which each pair of vertices is an edge. In other words, $E = \binom{V}{2}$. The complete graph with n vertices is denoted by K^n. A 1-factor of a graph G (of even order) is a set of independent edges spanning the vertices of G. A 1-factorization of G is a partition of the set of edges of G into 1-factors. A 1-factor is also called a *perfect matching*.

Given a complete graph with an even number of vertices, it is not difficult to show that there exists a 1-factorization. Let n be an even integer and consider the complete graph K^n. To find a factorization of K^n into 1-factors, select $n - 1$ vertices and place them on the vertices of a $n-1$ regular $n-1$-gon, and place the remaining vertex at the centre of the polygon. To get the first 1-factor, pick any vertex of the polygon and select the edge joining it to the vertex at the centre. For the remaining vertices, select the edges that are perpendicular to the line passing through the centre and the vertex already joined to it. For the remaining 1-factors, just select the $n - 2$ clockwise rotations of the first factor (see figure 2.1).

However, for the factorizations that we will consider, there will be restrictions on which edges can be selected for each factor. These restrictions will be defined in terms of the rows of a Hadamard matrix. A Hadamard matrix is a matrix with

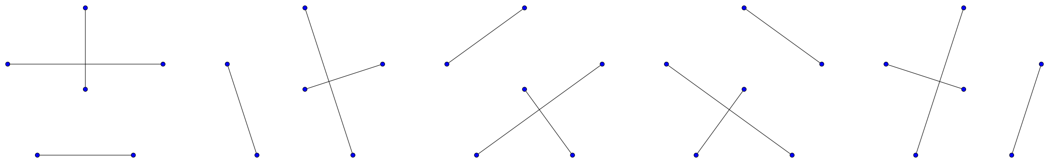


Figure 2.1: 1-factorization of K^6.

orthogonal rows (and orthogonal columns) whose entries are 1 or -1. Thus if H is a Hadamard matrix of order n, then $h_{ij} \in \{1, -1\}$ and

$$HH^* = nI_n$$

where I_n is the identity matrix of order n. It is easy to check that if $n > 2$ a Hadamard matrix can only exist if n is a multiple of 4. We will consider Hadamard matrices for which the first row consists of a vector all of whose entries are 1. Note that for any Hadamard matrix, it is always possible to transform this matrix to a matrix with the first row as desired by multiplying each column by the corresponding sign. We adopt this as the normalized version of a given Hadamard matrix. Hence, we will always assume that H is of the form

$$H = \left(\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ \hline \pm 1 & \pm 1 & \cdots & \pm 1 \\ \vdots & \vdots & \ddots & \vdots \\ \pm 1 & \pm 1 & \cdots & \pm 1 \end{array}\right)$$

In order to simplify our notation, we make the convention that the indices of the rows of a Hadamard matrix start from 0 so that the second row is $(h_{11}, h_{12}, \ldots, h_{1n})$ and subsequently up to the n-th row $(h_{(n-1)1}, h_{(n-1)2}, \ldots, h_{n-1,n})$:

$$H = \left(\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ \hline h_{11} & h_{12} & \cdots & h_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{(n-1)1} & h_{(n-1)2} & \cdots & h_{(n-1)n} \end{array}\right).$$

Given a Hadamard matrix H we want to find a 1-factorization $\{F_1, F_2, ..., F_{n-1}\}$ of K^n such that either each factor satisfies the restriction $R1$ below or each factor satisfies $R2$:

(R1) If an edge e belongs to the factor F_k, then the vertices incident to that edge must have *opposite* sign in the row k:

$$e = \{i, j\} \in F_k \implies h_{ki}h_{kj} < 0.$$

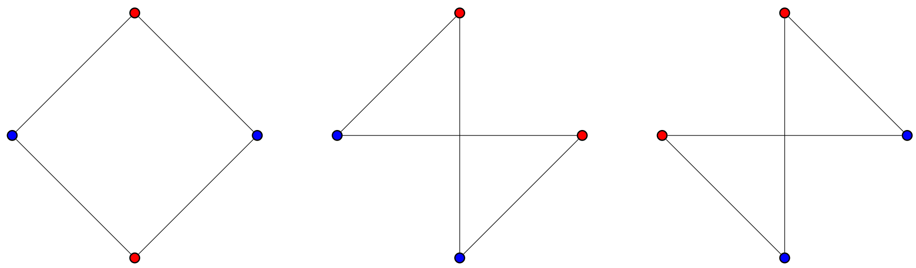
(R2) If an edge e belongs to the factor F_k, then the vertices incident to that edge must have *same* sign in the row k:

$$e = \{i, j\} \in F_k \implies h_{ki}h_{kj} > 0.$$

Let us illustrate the problem of finding a factorization satisfying restriction (R1) for the following example,

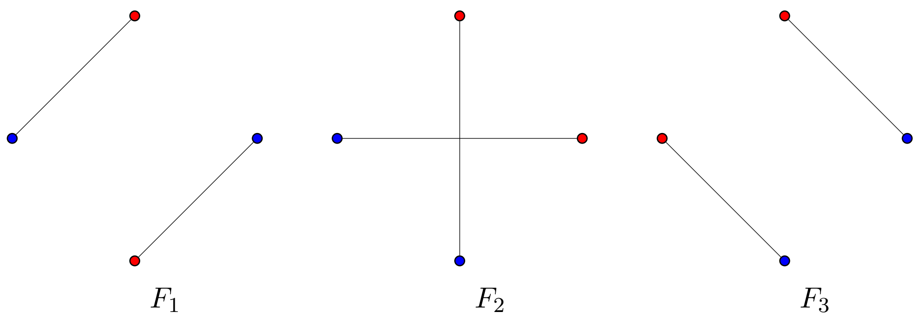
$$H = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

We want to find a 1-factorization $\{F_1, F_2, F_3\}$ of the complete graph on 4 vertices K^4 satisfying restriction (R1). For the row $(1, -1, 1, -1)$, there are 4 edges we could potentially select; $\{1, 2\}, \{2, 3\}, \{3, 4\}$, and $\{1, 4\}$. However, the only factors satisfying restriction (R1) are $\{\{1,2\},\{3,4\}\}$ and $\{\{1,4\},\{2,3\}\}$. We do the same analysis for the third row and we see that the only two possible factors are $\{\{1,3\},\{2,4\}\}$ and $\{\{2,4\},\{1,3\}\}$. Finally, for the fourth row we see that the only possible choices are $\{\{1,3\},\{2,4\}\}$ and $\{\{1,2\},\{3,4\}\}$.



Hence, from the 8 possible combinations of these pairs of edges a feasible 1-factorization of K^4 satisfying the requirements is

$$\begin{aligned} F_1 &= \{\{1,4\},\{2,3\}\}, \\ F_2 &= \{\{1,3\},\{2,4\}\}, \\ F_3 &= \{\{1,2\},\{3,4\}\}. \end{aligned}$$



For the general case of an arbitrary Hadamard matrix the problem seems to be far more complex than for the simple example: however we conjecture that it is always possible to find 1-factorizations satisfying the two different restrictions.

2.2 The Integer Program

An integer program is an optimization problem whose objective function and restrictions are all linear and such that the set of solutions is restricted to the integres. We can regard the problem as an integer program. We have a variable $x_{k,\{i,j\}}$ for each row k and pair of columns $\{i, j\}$. We want the variable to be either one or zero according as the edge $\{i, j\}$ belongs to the factor F_k. Hence, in the (R1) case, we want to find integer values $x_{k,\{i,j\}}$ such that

$$0 \leq x_{k,\{i,j\}} \leq \begin{cases} 0 & \text{if } h_{ki} = h_{kj} \\ 1 & \text{if } h_{ki} \neq h_{kj} \end{cases}, \tag{2.2.1}$$

and each edge $\{i, j\}$ in the complete graph K^n must appear only once in the factorization so we need

$$\sum_{k=1}^{n-1} x_{k,\{i,j\}} = 1, \tag{2.2.2}$$

and for each 1-factor we must select independent edges so for each $k \in \{1, \ldots, n-1\}$ and $j \in \{1, \ldots, n\}$

$$\sum_{i \neq j} x_{k,\{i,j\}} = 1. \tag{2.2.3}$$

If instead of (2.2.1) we ask the variables to satisfy

$$0 \leq x_{k,\{i,j\}} \leq \begin{cases} 0 & \text{if } h_{ki} \neq h_{kj} \\ 1 & \text{if } h_{ki} = h_{kj} \end{cases}, \tag{2.2.4}$$

then the solution to the integer program will be equivalent to finding a 1-factorization satisfying restriction (R2).

The linear relaxation of an integer program is the linear program with the same linear objective function and restrictions but where the solution is now allowed to be in the reals. The linear relaxation of our integer program is easily seen to be feasible and the Hadamard condition is just what is needed. Choose

$$x_{k,\{i,j\}} = \begin{cases} \frac{2}{n} & \text{if } h_{ki} \neq h_{kj} \\ 0 & \text{if } h_{ki} = h_{kj} \end{cases},$$

Since each row of H is orthogonal to the first row, for each k, i there are $n/2$ values of j such that the entries of the row k at the i-th and j-th columns have opposite sign and therefore $x_{k,\{i,j\}} = \frac{2}{n}$ for exactly $n/2$ values of j. Thus restriction (2.2.3) is satisfied.

On the other hand, each pair of columns of H is orthogonal. Hence, for each $\{i, j\}$ there are $n/2$ rows below the first one for which the entries in the i^{th} and j^{th} columns have opposite sign and therefore again $x_{k,\{i,j\}} = \frac{2}{n}$ for exactly $n/2$ values of k. Thus, restriction (2.2.2) is satisfied.

For the same sign case (restriction (R2)), we choose

$$x_{k,\{i,j\}} = \begin{cases} \frac{1}{n/2-1} & \text{if } h_{ki} = h_{kj} \\ 0 & \text{if } h_{ki} \neq h_{kj} \end{cases}.$$

In the two remaining sections of the chapter we show that the conjecture is true for some well-known classes of Hadamard matrices, Walsh and Paley matrices of certain sizes. As expected, finding factorizations for Walsh matrices will follow by a simple inductive argument. However, for Paley matrices, the construction of such factorizations will be considerably more sophisticated.

2.3 Factorizations for Walsh Matrices

The Walsh Matrices are constructed via an inductive process originally due to Sylvester. Given a Hadamard matrix H of order n we can construct a new Hadamard matrix of order $2n$ by defining the following matrix by blocks

$$\begin{pmatrix} H & H \\ H & -H \end{pmatrix}$$

Hence, we define

$$H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and we define inductively

$$H_m = \begin{pmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{pmatrix}$$

for each integer $m > 1$.

THEOREM 2.3.1. *Let $m > 1$ be an integer and $n = 2^m$. There exist 1-factorizations of K^n satisfying restrictions (R1) and (R2), respectively.*

Proof. The proof is by induction. We want to show first that there are such 1-factorizations for H_2

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

We already saw that there is a factorization satisfying restriction (R1) in our example. On the other hand it is easy to see that there is one and only one possible choice for a factorization satisfying restriction (R2).

Our inductive hypothesis states that there exist such factorizations satisfying (R1) and (R2) for H_{m-1}. As is common in this kind of induction we need both types of factorization for H_{m-1} to obtain each factorization of H_m but there is a strange additional issue to consider in one case.

Now

$$H_m = \left(\begin{array}{c|c} H_{m-1} & H_{m-1} \\ \hline H_{m-1} & -H_{m-1} \end{array}\right).$$

To find a factorization for H_m satisfying (R1) we decompose each of the H_{m-1} blocks in the top 2^{m-1} rows selecting edges of opposite sign. For the bottom 2^{m-1} rows we do as follows. The $2^{m-1}+1$ row consists of a block of 2^{m-1} positive entries and then a block of 2^{m-1} negative entries. We select the edges $\{i, i+2^{m-1}\}$ for all $i \in \{1, \ldots, 2^{m-1}\}$.

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{array}\right)$$

For the remaining rows, using our inductive hypothesis, we select edges of the form $\{i, j+2^{m-1}\}$ where the pair $\{i, j\}$ appears in a factorization of H_{m-1} satisfying restriction (R2). This gives us a factorization of H_m satisfying restriction (R1).

To find a factorization of H_m satisfying restriction (R2), we decompose each of the H_{m-1} blocks in the top 2^{m-1} rows selecting edges of the same sign. For the bottom 2^{m-1} rows we do as follows. The $2^{m-1}+1$ row consists of a block of 2^{m-1} positive entries and then a block of 2^{m-1} negative entries. We swap this row with any of the rows above, let's say the row 2^{m-1}, choosing the same edges on our new row 2^{m-1} that we already selected in the old row 2^{m-1}.

$$\left(\begin{array}{cccc|cccc} 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{array}\right)$$

To select edges in our new $2^{m-1}+1$ row, we select edges of the form $\{i, i+2^{m-1}\}$ which have the same sign since they come from the same entry in the matrix H_{m-1}.

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \end{array}\right)$$

For the remaining rows, we select edges of the form $\{i, j + 2^{m-1}\}$ where the pair $\{i, j\}$ appears in a factorization of H_{m-1} satisfying restriction (R1). This gives us a factorization of H_m satisfying restriction (R2). $\square$

When a paper on the subject of the material in this chapter was submitted to a journal, a referee suggested to us that the method used in the proof of Theorem 1.1 can easily be generalised. The proof shows that for $n \geq 2$, if there are 1-factorizations of H_n satisfying (R1) and (R2) respectively then there are 1-factorizations of $H_n \bigotimes H_1$ satisfying (R1) and (R2). A generalisation of this argument can be used to show that for any $m < n$, if there are 1-factorizations of H_n and H_m satisfying (R1) and (R2) then there are 1-factorizations of $H_n \bigotimes H_m$ satisfying (R1) and (R2).

2.4 Factorizations for Paley matrices and the finite field $\mathbb{Z}_p$

In this section we find 1-factorizations of the complete graph with restrictions defined in terms of matrices constructed using finite fields $\mathbb{Z}_p$ where p is a prime. This construction is due to Paley [27]. We shall use a slight variation of the usual Paley matrices. First we recall some basic terminology of number theory. A quadratic residue modulo n is an integer q that is congruent to a perfect square modulo n. In other words, q is a quadratic residue modulo n if there exists an integer x such that

$$q \equiv x^2 \pmod{n}.$$

Otherwise, q is called a quadratic nonresidue modulo n. Given a prime number p and an integer q the legendre symbol is define by

$$\left(\frac{q}{p}\right) = \begin{cases} 1 & \text{if } q \text{ is a quadratic residue mod } p \text{ and } q \not\equiv 0 \pmod{n} \\ -1 & \text{if } q \text{ is a quadratic residue mod } p \\ 0 & \text{if } q \equiv 0 \pmod{n} \end{cases}$$

There is a compact way to write the legendre symbol that follows from Euler's criterion (see [12] page 57):

$$\left(\frac{q}{p}\right) = a^{\frac{p-1}{2}}.$$

A basic property of the Legendre symbol is that it is a multiplicative function:

$$\left(\frac{qr}{p}\right) = \left(\frac{q}{p}\right)\left(\frac{r}{p}\right).$$

We now proceed to the construction of the Paley matrices. First we set

$$M = \left(\left(\frac{j-i}{p}\right)\right)_{i,j \in \mathbb{Z}_p}$$

where $\left(\frac{k}{p}\right)$ is the Legendre symbol. When $p \equiv 3 \pmod 4$, the Paley matrix H_p is defined to be

$$H_p = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1}^T & M \end{pmatrix} + I$$

where $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbb{Z}^p$ and I is the identity matrix of order $p+1$. To see that H_p is in fact a Hadamard matrix, notice that the row of M are just cyclic permutations of its first row. Hence, to show that the fisrt row of H_p is orthogonal to any other of its rows, it is enough to consider its product with the second row of H_p. This product is equal to

$$\left(\frac{1}{p}\right) + \left(\frac{2}{p}\right) + \cdots + \left(\frac{p-1}{p}\right) = 0,$$

since there are as many quadratic residues as nonresidues. To see that any other two rows are orthogonal, it is enough to consider the product of the second row by any of the rows below (again by the cyclicity of M). Let $r \not\equiv 0, 1 \pmod{p}$, then the product of the second row with the $r+1$-th row is equal to

$$1 + \left(\frac{r}{p}\right) + \left(\frac{-r}{p}\right) + \sum_{q=1}^{p-1} \left(\frac{q}{p}\right)\left(\frac{q+r}{p}\right). \tag{2.4.1}$$

Since $p \equiv 3 \pmod{4}$, it follows that

$$\left(\frac{-r}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{r}{p}\right) = (-1)^{\frac{p-1}{2}}\left(\frac{r}{p}\right) = -\left(\frac{r}{p}\right).$$

So (2.4.1) becomes

$$1 + \sum_{q=1}^{p-1} \left(\frac{q}{p}\right)\left(\frac{q+r}{p}\right).$$

On the other hand,

$$\left(\frac{q}{p}\right)\left(\frac{q+r}{p}\right) = \left(\frac{q(q+r)}{p}\right) = \left(\frac{1+lr}{p}\right),$$

where l is the reciprocal of $q \pmod{p}$. As q takes the values $1, 2, \ldots, p-2$, its reciprocal l also takes all values from 1 to $p-2$, since $(p-1)^2 = 1 \pmod{p}$. Hence,

$$1 + \sum_{q=1}^{p-1} \left(\frac{q}{p}\right)\left(\frac{q+r}{p}\right) = 1 + \sum_{l=1}^{p-1} \left(\frac{1+lr}{p}\right) = \sum_{m=1}^{p-1} \left(\frac{m}{p}\right) = 0.$$

This shows that the 2–th and $r+1$–th rows are othogonal. Hence H_p is a Hadamard matrix.

A primitive root of a prime number p is a integer x that has multiplicative order $p-1$. In other words, x is a primitive root of a prime p if x is a generator of the multiplicative cyclic group $(\mathbb{Z}_p^*, \cdot)$.

The main theorem of this section is the following in which a "near-primitive root" modulo p is just the square of a primitive root.

THEOREM 2.4.1. *Let p be a prime such that $p \equiv 3$ $(mod\ 4)$, then we can find a 1-factorization of the complete graph K^{p+1} satisfying restriction (R1) with respect to the Paley matrix of order $p+1$.*
If in addition we assume that 2 is a near-primitive root modulo p, then we can find a 1-factorization of the complete graph K^{p+1} satisfying restriction (R2).

It is natural to try to prove this theorem in the following way. Since the rows of M are just cyclic permutations of the first row it seems reasonable to find a 1-factor corresponding to the second row of H_p and then cycle it to obtain one 1-factors for the other rows just as in the example at the start of the chapter. This will work provided our 1-factor contains, within the matrix M, exactly one edge $\{i, j\}$ of each possible length: $i - j = \pm 1, \pm 2, \ldots, \pm(p-1)/2$. So we are led to consider the following problem which makes sense regardless of whether p is congruent to 1 or 3 modulo 4:

PROBLEM 2.4.2. *let p be any prime number, and K^{p+1} the complete graph with vertex set $\mathbb{Z}_p \cup \{\mathbf{c}\}$ where $\mathbf{c}$ is an additional point that we will call the centre. We adopt the convention that the length of any edge containing the centre is infinity, that the vertex 0 is a residue and that the centre $\mathbf{c}$ is a non-residue.*

(P1) Is there a 1-factor of K^{p+1} such that each edge selected is either incident to two quadratic residues or incident to two non-residues, and such that the lengths of the edges are all different to one another?

(P2) Is there a 1-factor of K^{p+1} such that each edge selected is incident to a quadratic residue and a non-residue, and such that the lengths of the edges are all different to one another?

There are two easy cases. The first one is when $p \equiv 1 \pmod 4$ and we want to join quadratic residues to quadratic residues, and non-quadratic residues to non-quadratic residues. The second one is when $p \equiv 3 \pmod 4$ and we want to join residues to non-residues. These two case are done using the following observation: -1 is a quadratic residue if and only if $p \equiv 1 \pmod 4$. Hence, when $p \equiv 1 \pmod 4$, we can select the edges of the form $\{r, -r\}$ where $r \in \{1, \ldots, \frac{p-1}{2}\}$, and by our previous observation, r and $-r$ are either both quadratic residues or both non-quadratic residues. The length of the edge $\{r, -r\}$ is $2r$ and all these lengths are clearly different to one another for $r \in \{1, \ldots, \frac{p-1}{2}\}$. We do exactly the same selection of edges when $p \equiv 3 \pmod 4$ but in this case we know we join quadratic residue to non-quadratic residues by our observation.

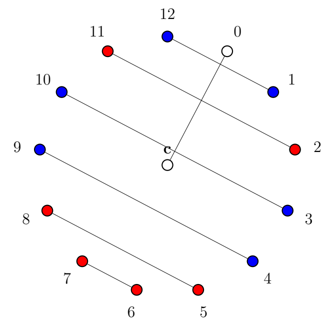


Figure 2.4: Joining quadratic residues to quadratic residues and non-quadratic residues to non quadratic residues for $p = 13$.

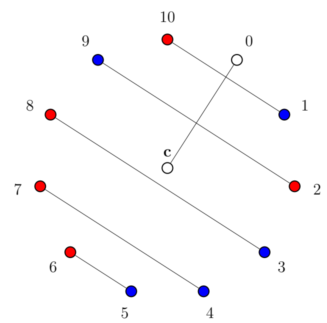


Figure 2.5: Joining quadratic residues to non-quadratic residues for $p = 11$.

Hence, we have the following theorems

THEOREM 2.4.3. *Let p be a prime such that $p \equiv 1$ (mod 4), then there exists a 1-factor of the complete graph with vertex set $\mathbb{Z}_p \cup \{c\}$ consisting of edges of all possible lengths matching residues to residues, and non-residues to non-residues.*

THEOREM 2.4.4. *Let p be a prime such that $p \equiv 3$ (mod 4), then there exists a 1-factor of the complete graph with vertex set $\mathbb{Z}_p \cup \{c\}$ consisting of edges of all possible lengths matching residues to non-residues.*

We now turn to the difficult cases. First, let p be a prime such that $p \equiv 1 \pmod 4$. In this case, we want to join residues to non-residues. Let x be a primitive root modulo p. We shall consider edges of the form $e_k = \{x^k, x^{k+1}\}$ where $k = 0, \dots, p-1$. Each of these edges joins a residue to a non-residue. The length of the edge is $x^{k+1} - x^k = x^k(x-1)$. These numbers are all different as k runs from 0 to $p-1$ but we wish to exclude the possibility that the edges that we choose include an opposite pair $\pm y$. The edges e_j and e_k have opposite lengths if

$$x^{k-j} = -1 = x^{(p-1)/2}.$$

Our aim will be to select $\frac{p-3}{2}$ which are different from one another and their negatives. These will join $p-3$ elements of $\mathbb{Z}_p^*$ and we wish to leave unjoined two elements: a residue r that we shall connect to 0 and a non-residue, n that we shall connect to the centre.

We therefore need to ensure that $\pm r$ is not one of the lengths that we have selected. Now if $r = x^k$ then we will not use the edge $\{x^k, x^{k+1}\}$ whose length is $x^k(x-1) = r(x-1)$ and this will indeed be r provided $x = 2$. Henceforth we assume this to be the case (which necessarily means that $p \equiv 5 \pmod 8$). It then doesn't matter which residue we choose for r so we take $r = 1 = 2^0$ and $n = -2 = 2^{(p+1)/2}$. We have used the vertex $r = 1$ and the length 1. We select the remaining edges to be

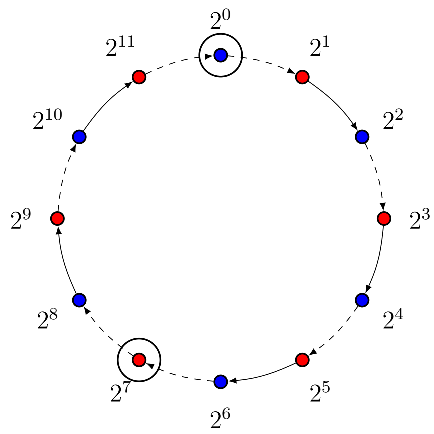


Figure 2.6: The figure shows the Cayley graph associated with $\mathbb{Z}_p^*$ for the generator 2 and the selection of edges for the case $p = 13$. The edges we select for the perfect matching are the solid lines joining quadratic residues to non-residues.

$$\begin{aligned}
e_1 = \{2, 4\} &= \{2^1, 2^2\} \\
e_3 = \{8, 16\} &= \left\{2^3, 2^4\right\} \\
&\vdots \\
e_{\frac{p-3}{2}} = \left\{2^{\frac{p-3}{2}}, -1\right\} &= \left\{2^{\frac{p-3}{2}}, 2^{\frac{p-1}{2}}\right\}
\end{aligned}$$

and

$$\begin{aligned}
e_{\frac{p+3}{2}} = \{-4, -8\} &= \left\{2^{\frac{p+3}{2}}, 2^{\frac{p+5}{2}}\right\} \\
e_{\frac{p+7}{2}} = \{-16, -32\} &= \left\{2^{\frac{p+7}{2}}, 2^{\frac{p+9}{2}}\right\} \\
&\vdots \\
e_{p-3} &= \left\{2^{p-3}, 2^{p-2}\right\}
\end{aligned}$$

For $k = 1, 3, \ldots, \frac{p-3}{2}$, the lenght of the edges e_k is equal to 2^k, and for $k = 1, 3, \ldots, \frac{p-5}{2}$ the length of the edge $e_{\frac{p+1}{2}+k}$ is equal to 2^{k+1} (see figure 2.6). Hence, we see that the lengths of the edges that we have selected are not equal nor equal to their negatives and we have not used the edges of length ± 1 which are e_0 and e_{p-2}. We are thus at liberty to join 1 to 0.

The second difficult case is that of a prime $p \equiv 3 \pmod 4$ and we want to join quadratic residues to quadratic residues and non-residues to non-residues. In this case we will assume that there a primitive root x modulo p such that $x^2 = 2$ (which necessarily implies that $p \equiv 7 \pmod 8$). Since $p \equiv 3 \pmod 4$ we know that -1 is not a quadratic residue.

We shall consider edges of the form $e_{2k} = \{x^{2k}, x^{2(k+1)}\}$ which join quadratic residues and edges of the form $e_{2k+(p-1)/2} = \{-x^{2k}, -x^{2(k+1)}\}$ $= \{x^{2k+(p-1)/2}, x^{2(k+1)+(p-1)/2}\}$ which join non-residues. The length of e_k is $x^{2(k+1)} - x^{2k} = x^{2k}(x^2 - 1) = x^{2k}$ and the length of $e_{2k+(p-1)/2}$ is $-x^{2k}$. These lengths are all different as k goes from 0 to $(p-3)/2$ since the negative of the length of e_{2k} is equal to the length of $e_{2k+(p-1)/2}$. The lengths of edges joining quadratic residues to quadratic residues are all different to one another and their negatives and the same is true for the edges joining non-residues to non-residues.

Our aim is to select $(p-3)/2$ which are different from one another and their negatives. These will join $p-3$ elements of $\mathbb{Z}_p^*$ and we wish to leave two elements alone: a quadratic residue r which we shall connect to 0 and a non-residue n which we shall connect to the center **c**.

We need to ensure that $\pm r$ is not one of the lengths that we have selected. We again take $r = 1$ and $n = -x^{p+3} = x^{2+(p-1)/2} = -2$. We select the remaining edges to be

$$
\begin{aligned}
e_2 = \{2, 4\} &= \{x^2, x^4\} \\
e_6 = \{8, 16\} &= \{x^6, x^8\} \\
\vdots \quad &\vdots \quad \vdots \\
e_{p-5} &= \{x^{p-5}, x^{p-3}\}
\end{aligned}
$$

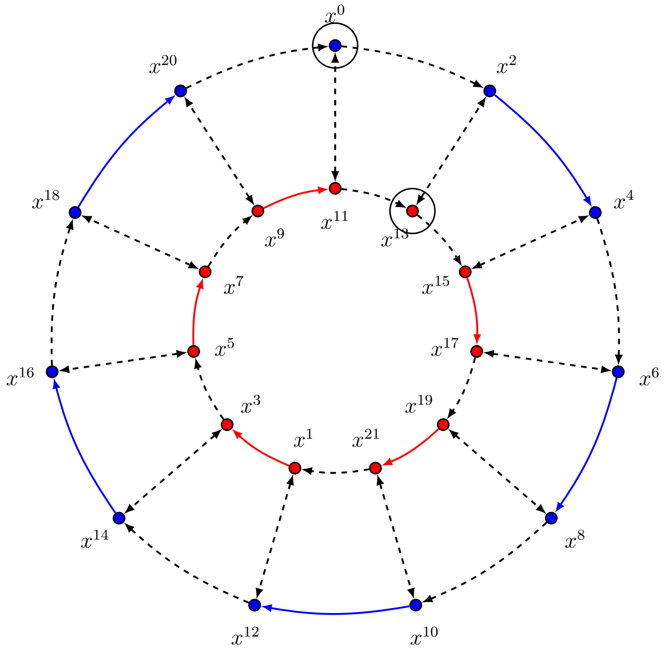


Figure 2.7: This figure shows Cayley graph associated with $\mathbb{Z}_p^*$ for the generating set $\{x^2, x^{(p-1)/2}\} = \{2, -1\}$ and the selection of edges incident to either two quadratic residues or two non-.residues, in solid blue and red lines respectively, for the case $p = 23$.

and

$$\begin{aligned}
e_{\frac{p-1}{2}+4} &= \left\{-x^4, -x^6\right\} \\
e_{\frac{p-1}{2}+8} &= \left\{-x^8, -x^{10}\right\} \\
\vdots \quad &\vdots \quad \vdots \\
e_{\frac{p-1}{2}+p-3} &= \left\{-x^{p-3}, -x^{p-1}\right\}.
\end{aligned}$$

For $k = 2, 6, \ldots, p-5$, the lenght of the edges e_k is equal to x^k, and for $k = 2, 6, \ldots, p-5$ the length of the edge $e_{\frac{p-1}{2}+k+2}$ is equal to x^{k+2} (see figure 2.7). Hence, the lengths of the edges that we have selected are not equal or equal to their negatives and we have not used the edges of length ± 1 which are e_0 and $e_{\frac{p-1}{2}}$.

To sum up, we have the following theorems.

THEOREM 2.4.5. *Let p be a prime such that $p \equiv 1 \pmod 4$ and 2 is a primitive root modulo p. Then there exists a 1-factor of the complete graph with vertex set $\mathbb{Z}_p \cup \{c\}$ consisting of edges of all possible lengths matching residues to non-residues.*

THEOREM 2.4.6. *Let p be a prime such that $p \equiv 3 \pmod 4$ and 2 is a near-primitive root of p. Then there exists a perfect matching of the complete graph with vertex set $\mathbb{Z}_p \cup \{c\}$ consisting of edges of all possible lengths matching residues to residues, and non-residues to non-residues.*

2.5 Further Remarks

Even though the proofs of Theorems 2.4.5 and 2.4.6 required an additional assumption on p, (concerning the number 2) we believe that they should be true in general. Theorem 2.4.5 is stated for primes p of the form $8k + 5$ for which 2 is a primitive root modulus p. It was pointed out to us by Peter Moree that under Generalized Riemann hypothesis there are infinitely many primes of this form and that these have a natural density which is a rational multiple of the Artin constant (see [22]). This is an example of a generalisation of Artin's conjecture asking for the density of primes p in an arithmetic progression such that an integer x is a primitive root modulo p.

On the other hand, Theorem 2.4.6 is stated for primes of the form $8k + 7$ for which there is a primitive root x modulo p such that $x^2 = 2$. In this case the question is whether there are infinitely many primes in an arithmetic progression for which a given integer t is a near primitive root. It was pointed out to us by Moree that this situation has not been worked out in the literature but that in our specific situation it would require no new ideas to do it.

To finish we remark that Problem 2.4.2 has a natural generalization.

PROBLEM 2.5.1. *Let $A \cup B$ be a partition of the cyclic group C_n where n is odd.*

Is it always possible to find a set of $(n-1)/2$ *edges* $\{x, y\}$ *with* $x, y \in C_n$*, whose* $n-1$ *lengths* $\pm(x-y)$ *include each non-zero element of* C_n *exactly once and so that each edge joins either two elements of* A *or two of* B.

We do not know of any counterexample to this problem: indeed we know of no counterexample even if we replace C_n with any finite group of odd order.

Chapter 3

Lipschitz Extensions

3.1 The Lipschitz Extension Problem

We will start by recalling basic definitions and useful notation. Recall that if $(\mathcal{M}, d_{\mathcal{M}})$ and $(\mathcal{N}, d_{\mathcal{N}})$ are metric spaces, we say that function $f : \mathcal{M} \to \mathcal{N}$ is Lipschitz if there exists a constant $K \in (0, \infty)$ such that

$$d_{\mathcal{N}}(f(x), f(y)) \leq K d_{\mathcal{M}}(x, y) \tag{3.1.1}$$

for all $x, y \in \mathcal{M}$. The smallest constant K for which (3.1.1) holds is called the Lipschitz constant of f and is denoted by $\|f\|_{\text{Lip}}$. Equivalently,

$$\|f\|_{Lip} = \sup \left\{ \frac{d_{\mathcal{N}}(f(x), f(y))}{d_{\mathcal{M}}(x, y)} : x, y \in \mathcal{M}, x \neq y \right\}. \tag{3.1.2}$$

Given a subset $\mathcal{A} \subset \mathcal{M}$ this can be can endowed with a metric structure inherited from the metric space $\mathcal{M}$. Given a Lipschitz function $f : \mathcal{A} \to \mathcal{N}$ we will say that $F : \mathcal{M} \to \mathcal{N}$ is an extension of f if F restricted to $\mathcal{A}$ is equal to f and F is Lipschitz. In this case, $\|F\|_{Lip} \geq \|f\|_{Lip}$.

Given metric spaces $(\mathcal{M}, d_{\mathcal{M}})$ and $(\mathcal{N}, d_{\mathcal{N}})$, the Lipschitz extension problem consists in determining whether an extension of Lipschitz function exists. Furthermore, if such extension exists, we are interested in estimating the multiplicative trade-off in the Lipschitz constant that has to be made in order to get such an extension. There are several variants of this problem; however, we can list them in three main categories: the general, the classical and the asymptotic variant.

Roughly speaking, the general variant asks whether given a subset $\mathcal{A} \subset \mathcal{M}$, there exists some constant $K \in (0, \infty)$ such that for any Lipschitz function $f : \mathcal{A} \to \mathcal{N}$ there exits an extension F on $\mathcal{M}$ such that

$$\|F\|_{Lip} \leq K \|f\|_{Lip} \tag{3.1.3}$$

The smallest constant K for which (3.1.3) holds is denote by $e(\mathcal{M}, \mathcal{N}; \mathcal{A})$.

The classical Lipschitz extension problem is a strengthening of the general one. It asks if there exists a constant K such that no matter the subset $\mathcal{A}$ of $\mathcal{M}$ nor the function $f : \mathcal{A} \to \mathcal{M}$ there exists always an extension F to the whole of $\mathcal{M}$ such that (3.1.3) holds. In this case, the smallest constant K such that (3.1.3) holds is denoted by $e(\mathcal{M}, \mathcal{N})$. Equivalently,

$$e(\mathcal{M}, \mathcal{N}) = \sup_{\mathcal{A} \subset \mathcal{M}} e(\mathcal{M}, \mathcal{N}; \mathcal{A}).$$

Finally when $e(\mathcal{M}, \mathcal{N}) = \infty$, the problem can be refined to understand the asymptotic behaviour of $e(\mathcal{M}, \mathcal{N}; \mathcal{A})$ as $\mathcal{A} \subset \mathcal{M}$ "grows" in some sense. There are two main ways of doing this. One way is by considering extensions from finite subsets to the whole metric space $\mathcal{M}$. Define the quantity $e_n(\mathcal{M}, \mathcal{N})$ as the supremum over all $e(\mathcal{M}, \mathcal{N}; A)$ such that $A \subset X$ and $|A| \leq n$. In this case, we are interested in understanding the asymptotic behaviour of $e_n(\mathcal{M}, \mathcal{N})$ as a function of $n \in \mathbb{N}$ when $n \to \infty$.

Another way is to consider extensions to finite subsets. In other words, consider an arbitrary subset Z of $\mathcal{M}$ and finite set A with $|A| \leq n$. Define $e^n(\mathcal{M}, \mathcal{N})$ to be the supremum of $e(Z \cup A, N; Z)$ over all possible choices of $A, Z \subset \mathcal{M}$ where $|A| \leq n$. In this case, we are interested in understanding the asymptotic behaviour of $e^n(\mathcal{M}, \mathcal{N})$ as a function of $n \in \mathbb{N}$ when $n \to \infty$.

3.2 Introduction

In their article, famous for their dimension reduction lemma, Johnson and Lindentrauss [14] studied extensions of Lipschitz maps from metric spaces into Hilbert space, proving that a Lipschitz map F defined on an n–point subset of a metric space to a Hilbert space H can be extended to a Lipschitz map $\tilde{F}$ defined on the whole metric space such that

$$||\tilde{F}||_{\text{Lip}} \leq K\sqrt{\log n}\, \|F\|_{\text{Lip}}$$

for some universal constant K. In the same article they posed the following question.

QUESTION 3.2.1. *Is there a non-linear analog of Maurey's Extension Theorem?*

Maurey's Extension Theorem relies on two important local properties of Banach spaces known as Type and Cotype.

DEFINITION 3.2.2. *For $p \leq 2$ and $q \geq 2$, the Type p constant and Cotype q constant of a Banach space $(X, \|\cdot\|_X)$, denoted by $T_p(X)$ and $C_q(X)$ respectively, are defined to be the smallest constants T and C, such that for any $x_1, ..., x_n \in X$,*

$$\mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|_X^p \leq T^p \sum_{i=1}^n \|x_i\|_X^p \quad \textit{and} \quad \sum_{i=1}^n \|x_i\|_X^q \leq C^q \mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|_X^q$$

where the expected value is taken over a uniform random choice of $\varepsilon \in \{-1, 1\}^n$. If $T_p(X)$ is finite, we say that X is a space of Type p; if $C_q(X)$ is finite, we say that the X is of Cotype q.

Maurey's extension theorem asserts the following.

THEOREM 3.2.3. *Given X a Banach spaces of Type 2, Y a Banach spaces of Cotype 2, Z a subspace of X and $u : Z \to Y$ a bounded linear operator, there exists a bounded linear extension $\tilde{u} : X \to Y$ such that*

$$\|\tilde{u}\|_{X\to Y} \leq T_2(X)C_2(Y)\|u\|_{Z\to Y}$$

Ball [3] not only answered Question 3.2.1 but also introduced appropriate metric versions of Type and Cotype, that imply an analog of Maurey's extension theorem for the general metric setting. He named this metric properties by Markov Type and Markov Cotype. Recall that a stochastic matrix A is a square matrix with non-negative entries whose rows add up to one.

DEFINITION 3.2.4. *The Markov Type p constant of a metric space $(\mathcal{M}, d_{\mathcal{M}})$, denoted by $M_p(\mathcal{M})$, is the smallest constant M such that for any $n \in \mathbb{N}$, $n \times n$ symmetric stochastic matrix $A = (a_{ij})$, $\alpha \in (0,1)$, and $x_1, ..., x_n \in \mathcal{M}$,*

$$(1-\alpha)\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} d_{\mathcal{M}}(x_i, x_j)^p \leq \alpha M^p \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} d_{\mathcal{M}}(x_i, x_j)^p.$$

where $C = (1-\alpha)(I - \alpha A)^{-1}$.

The Markov Cotype q constant of a vector space X, denoted by $N_q(X)$, is the smallest constant N such that for any $n \in \mathbb{N}$, any $n \times n$ symmetric stochastic matrix $A = (a_{ij})$, $\alpha \in (0,1)$, and $x_1, ..., x_n \in X$,

$$\alpha \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \left\| \sum_{r=1}^{n} c_{ir} x_r - \sum_{r=1}^{n} c_{jr} x_r \right\|_X^q \leq N^q (1-\alpha) \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \|x_i - x_j\|_X^q$$

where $C = (1-\alpha)(I - \alpha A)^{-1}$.

The name Markov Type and Cotype is explained by the following.

THEOREM 3.2.5 ([3]). *Let* $(\mathcal{M}, d)$ *be a metric space and* (M_n) *be a simple, symmetric Markov chain on* $\{1, \ldots, n\}$. *Then the following are equivalent*

(i) $\mathcal{M}$ *has Markov Type* p.

(ii) There is a constant K *so that if* $(M_k)_{k=1}^t$ *is a time-reversible stationary Markov chain on* $\{1, \ldots, t\}$, *running in steady state, and* $f : \{1, \ldots, t\} \to \mathcal{M}$,

$$\mathbb{E}d(f(M_t), f(M_0))^p \leq K^p t \mathbb{E}d(f(M_1), f(M_0))^p \tag{3.2.1}$$

(iii) There is a constant K *so that for any* $n \in \mathbb{N}$, *any* $n \times n$ *symmetric stochastic matrix* A, $t \in \mathbb{N}$, *and* $x_1, \ldots, x_n \in \mathcal{M}$ *we have that*

$$\sum_{ij} A_{ij}^t d(x_i, x_j)^p \leq K^p t \sum_{ij} a_{ij} d(x_i, x_j)^p.$$

In broad terms, the independence of the uniform choice of signs, that plays the main role in the definition of Type and Cotype, is replaced by the Markov condition in the metric space setting.

Remark 3.2.6. Ball [3] proposed a way to generalise the definition of Markov Cotype to metric spaces. This was later adapted by Mendel and Naor [20] to prove Ball's extension theorem when the target space is not necessarily a vector space. The Markov Cotype q constant of a metric space $(\mathcal{M}, d_{\mathcal{M}})$, denoted by $N_q(\mathcal{M})$, is the smallest constant N such that for any $n, t \in \mathbb{N}$, any $n \times n$ symmetric stochastic matrix $A = (a_{ij})$, and $x_1, ..., x_n \in \mathcal{M}$, there exists $y_1, ..., y_n \in \mathcal{M}$ such that

$$\sum_{i=1}^n d_{\mathcal{M}}(x_i, y_i)^q + \sum_{i=1}^n \sum_{j=1}^n a_{ij} d_{\mathcal{M}}(y_i, y_j)^q \leq N^q \sum_{i=1}^n \sum_{j=1}^n \frac{1}{t} \sum_{s=1}^t (A^s)_{ij} d_{\mathcal{M}}(x_i, x_j)^q$$

This definition of Markov Cotype 2 can be shown to agree with the one proposed by Ball [3]. A proof of this can be also found in [20]. Since we are just interested in the case when the target space is a vector space we will just focus on Ball's original definition for vector spaces.

With Markov Type and Cotype, Ball established a partial non-linear analogue of Maurey's theorem.

THEOREM 3.2.7 (Ball's Extension Theorem). *Let $\mathcal{M}$ be a a metric space of Markov Type* 2 *and Y a vector space of Markov Cotype* 2*, then*

$$e(\mathcal{M}, Y) \leq KM_2(\mathcal{M})N_2(Y)$$

where K is a universal constant.

Johnson and Lindentrauss [14] also ask the following long standing question.

PROBLEM 3.2.8. *For each $p \in (1,2)$ and $n \in \mathbb{N}$ there exist a constant K depending only on p such that if $\mathcal{A} \subset L_p$ and $u : \mathcal{A} \to \ell_2^n$ is Lipschitz then there exists a Lipschitz extension $\tilde{u}$ of u on L_p such that*

$$\|\tilde{u}\|_{Lip} \leq Kn^{\frac{1}{p}-\frac{1}{2}} \|u\|_{Lip}.$$

Problem 3.2.8 would be a non-linear version of Maurey's extension theorem for L_p spaces with $p \in (1,2)$. The linear version of this problem is due to König, Retherford and Tomczak-Jaegermann [17]. Their result is actually a strengthening of Maurey's extension theorem for finite-dimensional target spaces. This suggests that there could be a way to reformulate Problem 3.2.8 in more general terms.

THEOREM 3.2.9 ([17]). *If $q, p \in [1, \infty)$, X and Y are Banach spaces with $\dim Y = n$, and Z is a linear subspace of X, then for any linear operator $u : Z \to Y$ there exists a linear extension $\tilde{u} : X \to Y$ such that*

$$\|\tilde{u}\|_{X\to Y} \leq K_{p,q}T_p(X)C_q(Y)n^{\frac{1}{p}-\frac{1}{q}} \|u\|_{Z\to Y}$$

where the constant $K_{p,q}$ depends only on p and q.

Naor and Rabani [24] proposed the following problem as a nonlinear analog of theorem 3.2.9.

PROBLEM 3.2.10. *Suppose that $1 \leq p \leq q < \infty$ and $(\mathcal{M}, d_{\mathcal{M}})$ is a metric space with Markov Type p. Suppose also that $n \in \mathbb{N}$ and that $(Y, \|\cdot\|_Y)$ is an n-dimensional normed space. Is it true that there exists a constant*

$$K = K(M_p(\mathcal{M}), N_q(Y), p, q)$$

which may depend only on $M_p(\mathcal{M}), N_q(Y), p, q$ such that

$$e(X, Z) \leq K n^{\frac{1}{p} - \frac{1}{q}}?$$

An elegant proof of Theorem 3.2.9 by Pisier [28] gives us a hint on how problem 3.2.10 could be tackled. In his proof, Pisier uses the following Type 2 property for spaces with Type p where $p \in (1, 2)$.

PROPOSITION 3.2.11 ([28]). *Let X be space of Type $p \in (1, 2)$ and $A = (a_{ij}) \in \mathcal{M}_{n\times m}(\mathbb{R})$ be a contraction, i.e.*

$$\|A\|_{\ell^2_m \to \ell^2_n} \leq 1.$$

For any finite sequence $x_1, ..., x_m \in X$,

$$\mathbb{E}\left\| \sum_{i=1}^{n} \varepsilon_i \sum_{j=1}^{m} a_{ij} x_j \right\|_X^2 \leq K_p n^{\frac{2}{p}-1} \sum_{j=1}^{m} \|x_j\|_X^2$$

where K_p is a constant depending only on p.

The outline of the chapter is the following. In the first section, we give a simple proof of Proposition 3.2.11 for the case when $X = L_p$ for some $p \in [1, 2]$ using an interpolation argument. We explain how this proposition is used to obtain Theorem 3.2.9 for L_p spaces. In the next section, We propose a possible metric analogue of Proposition 3.2.11:

CONJECTURE. *Let $1 < p < 2$, $x_1, \ldots, x_n \in L_p$ and A an $n \times n$ symmetric stochastic matrix. Then, for any positive integer t we have that*

$$\sum_{ij} A^t_{ij} \left\|x_i - x_j\right\|^2_{L_p} \lesssim_p (\log n)^{\frac{2}{p}-1} t \sum_{ij} a_{ij} \left\|x_i - x_j\right\|^2_{L_p}.$$

We prove Conjecture 3.3.2 for $p = 1$ (actually, for an arbitrary metric space) and obtain a weaker version of this conjecture for general p.

THEOREM. *Suppose that $1 < p < 2$. Let $x_1, \ldots, x_n$ be a finite sequence of n vectors in L_p and A an $n \times n$ symmetric stochastic matrix. Then, for all $t \in \mathbb{N}$ and $q < p$ we have that*

$$\sum_{ij} A^t_{ij} \left\|x_i - x_j\right\|^q_{L_p} \lesssim_{p,q} (\log n)^{q/p-q/2} t \sum_{ij} a_{ij} \left\|x_i - x_j\right\|^q_{L_p}$$

After that, we show the validity of Conjecture 3.3.2 for what we believe should be the worst possible case: when A is the transition matrix of the standard random walk on a hypercube (actually we prove it for any transition matrix for which the transition probabilities depend only on the Hamming distance between vertices). Finally, we give some results concerning the extension of Lipschitz maps defined on finitely many points in L_p spaces with $p \in [1, 2]$ which later were found independently by Mendel and Naor [21]. Our main theorem of the last section is the following:

THEOREM. *Let $p \in (1, 2)$ and Y a vector space with Markov Cotype 2. Then*

$$e^n(L_p, Y) \lesssim_p (\log n)^{\frac{1}{p}-\frac{1}{2}} N_2(Y).$$

3.2.1 Proof of Proposition 3.2.11 for L_p

In the following paragraphs we give a simple proof of proposition 3.2.11 for the case when $X = L_p([0,1])$ for some $p \in [1,2]$ and we explain how this is applied to get 3.2.9 for the particular case when $X = L_p$, for $p \in [1,2]$ and $Y = \ell_n^2$. We introduce some standard notation that we will be using throughout the chapter. We write $a \lesssim b$ whenever there is a universal constant K such that $a \le Kb$. Similarly, we write $a \lesssim_p b$ if there is a constant K depending on p such that $a \le Kb$. Finally, we write $a \asymp b$ if $a \lesssim b$ and $b \lesssim a$; and $a \asymp_p b$ if $a \lesssim_p b$ and $b \lesssim_p a$.

Remark 3.2.12. If X is a Banach space, $x_1, x_2, \dots, x_n$ is a finite sequence of elements of X and $\theta_1, ..., \theta_n$ are real numbers then, by a simple application of Cauchy-Schwarz inequality, it follows that

$$\left\| \sum_{i=1}^n \theta_i x_i \right\|^2 \le \left(\sum_{i=1}^n \theta_i^2 \right) \sum_{i=1}^n \|x_i\|^2$$

Hence if A is a $n \times m$ real matrix such that $\|A\|_{\ell_m^2 \to \ell_n^2} \le 1$, then

$$\begin{aligned}
\mathbb{E} \left\| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j \right\|^2 &\le \mathbb{E} \sum_{j=1}^m \left(\sum_{i=1}^n a_{ij} \varepsilon_i \right)^2 \sum_{j=1}^m \|x_j\|^2 \\
&= \mathbb{E} \left\| A^\top \varepsilon \right\|_{\ell_m^2}^2 \sum_{j=1}^m \|x_j\|^2
\end{aligned}$$

Since $A^\top$ is a contraction, we have that

$$\left\| A^\top \varepsilon \right\|_{\ell_n^2} \le \|\varepsilon\|_{\ell_n^2} = n.$$

Hence,

$$\mathbb{E} \left\| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j \right\|^2 \le n \sum_{j=1}^m \|x_j\|^2$$

Proof of lemma 3.2.11 for L_p. By Khintchine inequality [16], for all $t \in [0,1]$,

$$\mathbb{E}_\varepsilon \left| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j(t) \right|^p \asymp_p \left(\sum_{i=1}^n \left(\sum_{j=1}^m a_{ij} x_j(t) \right)^2 \right)^{\frac{p}{2}} \tag{3.2.2}$$

Integrating the left hand side of (3.2.2) with respect to t we get

$$\begin{aligned} \int \mathbb{E}_\varepsilon \left| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j(t) \right|^p dt &= \mathbb{E}_\varepsilon \int \left| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j(t) \right|^p dt \\ &= \mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j \right\|_{L_p}^p dt \end{aligned}$$

Hence,

$$\begin{aligned} \left(\mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j \right\|_{L_p}^p \right)^{\frac{1}{p}} &\asymp_p \left(\int \left(\sum_{i=1}^n \left(\sum_{j=1}^m a_{ij} x_j(t) \right)^2 \right)^{\frac{p}{2}} dt \right)^{\frac{1}{p}} \\ &= \left(\int \|Ax(t)\|_{\ell_n^2}^p \, dt \right)^{\frac{1}{p}} \end{aligned}$$

where

$$x(t) = (x_1(t), \ldots, x_m(t)) \tag{3.2.3}$$

for all t. By Kahane's inequality [15] we may replace the exponent p by 2 on the left hand side of the above inequality to get

$$\left(\mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i \sum_{j=1}^m a_{ij} x_j \right\|_{L_p}^2 \right)^{\frac{1}{2}} \asymp_p \left(\int \|Ax(t)\|_{\ell_n^2}^p \, dt \right)^{\frac{1}{p}} \tag{3.2.4}$$

The vector x defined as in (3.2.3) can be consider as a function in the spaces $L_{2,p}([m] \times [0,1])$, the space of functions defined on the product space $[m] \times [0,1]$, where $[m] = \{1, \ldots, m\}$, endowed with the norm

$$\|x\|_{L_{(2,p)}} = \left(\int \left(\sum_{j=1}^{m} x_j^2(t) \right)^{\frac{p}{2}} dt \right)^{\frac{1}{p}} = \left(\int \|x(t)\|_{\ell_m^2}^p \, dt \right)^{\frac{1}{p}}.$$

Define the operator $\mathcal{A} : L_{(p,2)}([0,1] \times [m]) \to L_{(2,p)}([n] \times [0,1])$ by

$$\mathcal{A}x = Ax(t) = \left(\sum_{j=1}^{m} A_{ij} x_j(t) \right)_{i=1}^{n}$$

for all $x \in L_{(p,2)}([n] \times [0,1])$. By (3.2.4) it follows that

$$\left(\mathbb{E}_\varepsilon \left\| \sum_{i=1}^{n} \varepsilon_i \sum_{j=1}^{m} a_{ij} x_j \right\|_{L_p}^2 \right)^{\frac{1}{2}} \asymp_p \|\mathcal{A}x\|_{L_{(2,p)}}$$

So our problem is to show that

$$\|\mathcal{A}\|_{L_{(p,2)} \to L_{(2,p)}} \precsim_p n^{\frac{1}{p} - \frac{1}{2}},$$

Since A is a contraction, it is easy to see that

$$\|\mathcal{A}\|_{L_{(2,2)} \to L_{(2,2)}} \leq 1.$$

On the other hand, by remark 3.2.12, taking $X = L_1$ it follows that

$$\|\mathcal{A}\|_{L_{(2,1)} \to L_{(1,2)}} \leq \sqrt{n}.$$

Thus, applying Riesz-Thorin interpolation theorem for L_p spaces with mixed norms (see section 7 theorem 2 in [9]) it follows that

$$\|\mathcal{A}\|_{L_{(2,p)} \to L_{(p,2)}} \precsim_p n^{\frac{\theta}{2}} \tag{3.2.5}$$

for θ such that $\frac{1}{p} = \frac{\theta}{1} + \frac{(1-\theta)}{2}$ and hence $\theta = \frac{2}{p} - 1$. $\square$

To see how this Proposition 3.2.11 is applied to obtain Theorem 3.2.9, we need to state the following classic criterion for extension of linear operators between vector spaces.

THEOREM 3.2.13 ([28]). *Let X and Y be Banach spaces, $Z \subset X$ a linear subspace of X and $u : Z \to Y$ a bounded linear operator. Then there exists an extension $\tilde{u} : X \to Y$ of u satisfying $\|\tilde{u}\|_{X \to Y} \leq K$ if and only if for any finite sequence $z_1, z_2, \ldots z_m$ in Z and any $k \times m$ contraction A,*

$$\sum_{i=1}^{k} \left\| \sum_{j=1}^{m} a_{ij} u(z_j) \right\|_Y^2 \leq K \sum_{j=1}^{m} \|z_j\|_X^2$$

Following an observation by Maurey in his study of factorization and extension of linear operators, if a space has Type 2 then the Type 2 property is satisfied if we replaces the Bernoulli random variables by Gaussian random variables.

Given the fact that Kahane's inequality is true for Gaussian sums, the proof of Proposition 3.2.11 is valid if we replace Bernoulli random variables by Gaussian random variable as well. The main advantage of replacing the Bernoulli random signs by Gaussian random variables is that the latter is invariant under orthogonal transformations which will play an important role in the sequel. From now on we denote by $(g_i)_{i=1}^{\infty}$ a sequence of independent and identically distributed standard gaussian random variables. Recall that if $g_1, \ldots, g_n$ are independent and identically distributed standard gaussian random variables, their joint density is given by

$$(2\pi)^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2}\|x\|_2^2\right)$$

where $x = (x_1, \ldots, x_n)$. Hence, for any measurable set $A \subset \mathbb{R}^n$,

$$P((g_1, \ldots, g_n) \in A) = \int_A (2\pi)^{-n/2} \exp\left(-\frac{1}{2}\|x\|_2^2\right) dx.$$

Hence, for any $n \times n$ orthogonal matrix Q the vectors Qg and gQ will have the same density as g, since $\|x\|_2^2 = \|Qx\|_2^2 = \|xQ\|_2^2$ for all $x \in \mathbb{R}^n$.

LEMMA 3.2.14 ([28]). *Let X be a Banach space and $A = (a_{ij}) \in \mathcal{M}_{n\times m}(\mathbb{R})$ be a contraction. For any finite sequence $x_1, ..., x_m \in X$,*

$$\mathbb{E}\left\|\sum_{i=1}^{n} g_i \sum_{j=1}^{m} a_{ij}x_j\right\|_X^2 \leq \left\|\sum_{j=1}^{m} g_j x_j\right\|_X^2 .$$

Proof. We assume that $n = m$. Let $\mathcal{A}$ be the set of all $n \times n$ contractions and define the function $F : \mathcal{A} \to \mathbb{R}$ by

$$F(A) = \mathbb{E}\left\|\sum_{i=1}^{n} g_i \sum_{i=1}^{n} a_{ij}x_j\right\|_X^2$$

It is easy to see that F is convex. Hence, F attains its maximum on a extreme point of $\mathcal{A}$; hence, on an orthogonal matrix Q. Since Q is orthogonal and the Gaussian distribution is invariant under orthogonal transformation we have that $(\sum_{i=1}^{n} g_i q_{ij})_j$ has the same distribution as (g_i) and hence

$$F(Q) = \mathbb{E}\left\|\sum_{i=1}^{n} g_i \sum_{j=1}^{n} q_{ij}x_j\right\|_X^2 = \mathbb{E}\left\|\sum_{j=1}^{n} (gQ)_j x_j\right\|_X^2 = \mathbb{E}\left\|\sum_{j=1}^{n} g_j x_j\right\|_X^2 .$$

□

The following lemma plays a crucial role in the proof of Theorem 3.2.9.

LEMMA 3.2.15. *Let Z be a vector space, $u : Z \to \ell_2^n$ be a linear operator and $z_1, z_2, \ldots, z_m \in Z$. Then there exist a projection $P \in \mathcal{M}_{m\times m}$ and a contraction*

$Q \in \mathcal{M}_{m\times n}$ *such that if we set*

$$\begin{aligned}\tilde{z}_i &= \sum_{j=1}^{m} P_{ji} z_j \\ w_k &= \sum_{j=1}^{m} Q_{jk} z_j\end{aligned}$$

for all $i \in \{1, \dots, m\}$ *and* $k \in \{1, \dots, n\}$ *then*

$$\tilde{z}_i = \sum_{k=1}^{n} Q_{ik} w_k$$

and $u(\tilde{z}_i) = u(z_i)$ *for all* $i \in \{1, \dots, m\}$.

Proof. Let M be the $n \times m$ matrix defined by

$$M = \left(\begin{array}{c|c|c|c} u(z_1) & u(z_2) & \cdots & u(z_m) \end{array} \right)$$

and let $v_1, \dots, v_k$ be an orthonormal basis for the range of M^T. Note that $k = rank(M^T) = rank(M) \le n$. Let Q be the following matrix defined by columns

$$Q = \left(\begin{array}{c|c|c|c|c} v_1 & v_2 & \cdots & v_k & 0_{n-k\times n} \end{array} \right)$$

and let $P = QQ^T$. Hence, consider as a map from $\mathbb{R}^m$ to $\mathbb{R}^m$

$$P = QQ^\top = \sum_{i=1}^{n} v_i \otimes v_i$$

which is just the orthogonal projection onto the range of M^T. Finally, let

$$\begin{aligned}\tilde{z}_i &= \sum_{j=1}^{m} P_{ji} z_j \\ w_k &= \sum_{j=1}^{m} Q_{jk} z_j\end{aligned}$$

for all $j \in \{1, \ldots, m\}$ and $k \in \{1, \ldots, n\}$. Then,

$$\tilde{z}_i = \sum_{i=1}^{m} P_{ji} z_j = \sum_{i=1}^{m} (QQ^T)_{ji} z_j = \sum_{k=1}^{n} Q_{ik} w_k$$

and

$$u(\tilde{z}_i) = \sum_{j=1}^{m} P_{ji} u(z_j) = MPe_i = Me_i = u(z_i) \tag{3.2.6}$$

for all $i \in \{1, \ldots, m\}$. $\square$

In the proof of Lemma 3.2.15, we used for the first time the fact that our linear map has a finite dimensional target. This lemma allows us to construct n vectors $(w_i)_{i=1}^n$ which will serve as an anchor to be able to apply Proposition 3.2.11.

THEOREM 3.2.16 ([28]). *Let $p \in (1, 2)$. Let Z a subspace of L_p and $u : Z \to \ell_2^n$ a bounded linear operator. Then, there exists bounded linear extension $\tilde{u} : L_p \to \ell_2^n$ of u such that*

$$||\tilde{u}||_{L_p \to \ell_2^n} \precsim_p n^{\frac{1}{p} - \frac{1}{2}} ||u||_{Z \to \ell_2^n}$$

Proof. We will use the extension criterion from Theorem 3.2.13. Let $z_1, \ldots, z_m \in Z$ and let A be a $l \times m$ contraction. Let P, Q, $\tilde{z}_1, \ldots, \tilde{z}_m$ and $w_1, \ldots, w_n$ as in Lemma 3.2.15. Since $u(z_i) = u(\tilde{z}_i)$ for all $i \in \{1, 2, \ldots, m\}$ we have that

$$\sum_{i=1}^{l} \left\| \sum_{j=1}^{m} a_{ij} u(z_j) \right\|_{\ell_2^n}^2 = \sum_{i=1}^{l} \left\| \sum_{j=1}^{m} a_{ij} u(\tilde{z}_j) \right\|_{\ell_2^n}^2 \tag{3.2.7}$$

Since ℓ_2^n is a Hilbert space it is easy to see that

$$\sum_{i=1}^{l} \left\| \sum_{j=1}^{m} a_{ij} u(\tilde{z}_j) \right\|_{\ell_2^n}^2 = \left\| \sum_{i=1}^{l} g_i \sum_{j=1}^{m} a_{ij} u(\tilde{z}_j) \right\|_{\ell_2^n}^2 \tag{3.2.8}$$

Using the fact that u is a bounded linear operator we have that (3.2.8) is at most

$$\|u\| \left\| \sum_{i=1}^{l} g_i \sum_{j=1}^{m} a_{ij} \tilde{z}_j \right\|_{L_p}^2 \tag{3.2.9}$$

On the other hand, by Lemma 3.2.14

$$\left\| \sum_{i=1}^{l} g_i \sum_{j=1}^{m} a_{ij} \tilde{z}_j \right\|_{\ell_2^n}^2 \leq \left\| \sum_{j=1}^{m} g_j \tilde{z}_j \right\|_{L_p}^2 \tag{3.2.10}$$

and by Lemma 3.2.14 and 3.2.15

$$\left\| \sum_{j=1}^{m} g_j \tilde{z}_j \right\|_{L_p}^2 \leq \left\| \sum_{k=1}^{n} g_k w_k \right\|_{L_p}^2 \tag{3.2.11}$$

Apply Proposition 3.2.11 we obtain

$$\left\| \sum_{k=1}^{n} g_k w_k \right\|_{L_p}^2 = \left\| \sum_{k=1}^{n} g_k \sum_{j=1}^{m} Q_{jk} z_j \right\|_{L_p}^2 \precsim_p n^{\frac{2}{p}-1} \sum_{j=1}^{m} \|z_j\|_{L_p}^2. \tag{3.2.12}$$

Hence,

$$\sum_{i=1}^{l} \left\| \sum_{j=1}^{m} a_{ij} u(z_j) \right\|_{\ell_2^n}^2 \precsim_p \|u\|_{Z \to \ell_2^n} n^{\frac{2}{p}-1} \sum_{j=1}^{m} \|z_j\|_{L_p}^2. \tag{3.2.13}$$

Therefore, the desire result follows from Theorem 3.2.13. □

3.3 Asymptotic Markov Type 2 for L_p with $p \in (1, 2)$

Proposition 3.2.11 is one of the main ingredients in the proof of the extension Theorem 3.2.9. In order to tackle Problem 3.2.8, one might expect to find a similar property in the non-linear setting: a Markov Type 2 property for L_p spaces with $p \in [1, 2)$. A first step in this direction is the following theorem which is just an analogue of Remark 3.2.12.

THEOREM 3.3.1. *Let (M,d) be any metric space, $x_1,\dots,x_n \in M$ and A an $n\times n$ symmetric stochastic matrix. Then, for all $t\in\mathbb{N}$ we have that*

$$\sum_{ij} A^t_{ij} d(x_i,x_j)^2 \leq 4(\log n)t\sum_{ij} a_{ij}d(x_i,x_j)^2.$$

Proof of theorem 3.3.1. First note that any n-points on a metric space can be embedded isometrically into ℓ_∞^n, i.e. there exists a map $\varphi:\{x_1,\dots,x_n\}\to\ell_\infty^n$ such that

$$\|\phi(x_i)-\phi(x_j)\|_\infty = d(x_i,x_j) \tag{3.3.1}$$

for all $i,j\in\{1,\dots,n\}$. For instance, one can use the following embedding defined by

$$f(x_i)=\sum_{k=1}^n d(x_i,x_k)e_k$$

for all $i\in\{1,\dots,n\}$. Hence, triangle inequality implies

$$d(x_i,x_j)\geq \sup\{|d(x_i,x_k)-d(x_j,x_k)|\} = \|f(x_i)-f(x_j)\|_\infty$$

and

$$\|f(x_i)-f(x_j)\|_\infty \geq |d(x_i,x_i)-d(x_j,x_i)| = d(x_i,x_j) \tag{3.3.2}$$

for all $i,j\in\{1,\dots,n\}$. On the other hand, by equivalence of norms $M_2(\ell_\infty^n)\leq n^{\frac{1}{p}}M_2(\ell_p^n)$ for all $p>0$ (see Definition 3.2.4). Since $M_2(\ell_p^n)\leq 4\sqrt{p}$ (see [23] Theorem 1.2),

$$M_2(\ell_\infty^n)\leq 4n^{\frac{1}{p}}\sqrt{p}.$$

In fact, if we take $p=\log n$ it follows that

$$M_2(\ell_\infty^n)\leq 4\sqrt{\log n}.$$

Hence, by Proposition 3.2.5 we have that

$$\sum A_{ij} d(x_i, x_j)^2 = \sum A_{ij} \|f(x_i) - f(x_j)\|_\infty^2$$
$$\leq 4(\log n) t \sum_{ij} a_{ij} \|f(x_i) - f(x_j)\|_\infty^2 = (\log n) t \sum_{ij} a_{ij} d(x_i, x_j)^2 \quad (3.3.3)$$

□

When the metric space $\mathcal{M}$ is replaced by L_2 it is known that the Markov Type 2 constant is 1 and therefore we don't need any asymptotic factor depending on the number of points. Hence, we believe the correct metric property analogue to Proposition 3.2.11 should be the one stated in the following conjecture.

CONJECTURE 3.3.2. *Let* $1 < p < 2$, $x_1, \dots, x_n \in L_p$ *and* A *an* $n \times n$ *symmetric stochastic matrix. Then, for all* $t \in \mathbb{N}$ *we have that*

$$\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^2 \precsim_p (\log n)^{\frac{2}{p}-1} t \sum_{ij} a_{ij} \|x_i - x_j\|_{L_p}^2$$

The most general result that we have in this direction is the following new theorem.

THEOREM 3.3.3. *Suppose that* $1 < p < 2$. *Let* $x_1, \dots, x_n$ *be a finite sequence of* n *vectors in* L_p *and* A *an* $n \times n$ *symmetric stochastic matrix. Then, for all* $t \in \mathbb{N}$ *and* $q < p$ *we have that*

$$\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^q \precsim_{p,q} (\log n)^{q/p - q/2} t \sum_{ij} a_{ij} \|x_i - x_j\|_{L_p}^q$$

In order to prove theorem 3.3.3, we will need to introduce a construction of a p-stable process with parameter the L_p-norm due to Marcus and Pisier [19].

3.4 A p-stable processes with parameter the L_p norm, $p \in [1,2]$

A complete exposition of properties of p-stable process and the results that are stated here can be found in [19]. Let $(S, \mathcal{S}, \mu)$ be a finite measure spaces. Without loss of generality, we will assume that μ is a probability measure on S. We will consider three sequences of random variables independent of each other. First, let $\{U_j\}_{j=1}^{\infty}$ be a sequence of i.i.d S-valued random variables with $P(U_j \in A) = \mu(A)$ for all $A \in \mathcal{S}$. Second, Let $\{\theta_j\}_{j=1}^{\infty}$ sequences of i.i.d. exponential random variables with parameter one and define $\Gamma_j = \theta_1 + \cdots + \theta_j$. Therefore Γ_j is distributed as a gamma random variable with parameters k and 1. Finally, let $\{g_j\}_{j=1}^{\infty}$ a sequence of i.i.d. standard Gaussian random variables.

We may assume also that the sequences $\{U_j\}$ and $\{\theta_j\}$ are defined on the same probability spaces $(\Omega, \Sigma, \mathbb{P})$, and $\{g_j\}_{j=1}^{\infty}$ is defined on a different probability space $(\Omega', \Sigma', \mathbb{P}')$.
By proposition 1.5 in [19], there exists a constant C_p depending only on p such that the stochastic process defined by

$$T(x) = C_p \sum_{j=1}^{\infty} (\Gamma_j)^{-\frac{1}{p}} g_j x(U_j) \tag{3.4.1}$$

for each $x \in L_p$ is a p-stable process and

$$\mathbb{E} e^{iT(x)} = e^{-\|x\|_{L_p}^p} \tag{3.4.2}$$

for all $x \in L_p$. Let us denote by $T_{(\omega,\omega')}(x)$ a realization of the random variable $T(x)$. Notice also that, by varying $x \in L_p$, this process can be though as a distribution on $\mathcal{L}(L_p, \mathbb{R})$, the space of linear functionals on L_p. By fixing $\omega \in \Omega$ notice that

$T_\omega := T_{(\omega,\cdot)}x$ is actually a normal random variable with variance $\|T_\omega(x)\|_{L_2}^2$ and therefore this gives us a distribution on the linear operators from L_p to L_2 such that

$$\mathbb{E}_{\mathbb{P}'}\left[e^{iT_\omega(x)}\right] = e^{-\|T_\omega(x)\|_{L_2}^2} \tag{3.4.3}$$

for all $x \in L_p$ Taking expectation with respect to $\mathbb{P}$ and using (3.4.2) and (3.4.3), we get that for all $\lambda \geq 0$

$$\mathbb{E}_{\mathbb{P}}\left[e^{-\lambda\|T_\omega(x)\|_{L_2}^2}\right] = e^{-\|x\|_{L_p}^p \lambda^{\frac{p}{2}}}. \tag{3.4.4}$$

Since $T(x)$ is a p-stable random variable with parameter $\|x\|_p$ it follows that for every $0 < r < p$ there exists a constant δ_{rp} depending on r and p only such that

$$(\mathbb{E}|T(x)|^r)^{1/r} = \delta_{rp}\|x\|_p. \tag{3.4.5}$$

By (3.4.4) we have that

$$\mathbb{E}_{\mathbb{P}}\left[e^{-\lambda\left(\frac{\|T_\omega(x)\|_{L_2}^2}{\|x\|_{L_p}^2}\right)}\right] = e^{-\lambda^{\frac{p}{2}}}. \tag{3.4.6}$$

A simple application of Chebysev's inequality yields the following large deviation bound

$$\mathbb{P}\left(\left\{\omega \in \Omega : \|T_\omega(x)\|_{L_2} < \varepsilon \|x\|_{L_p}\right\}\right) \leq e^{-\lambda^{\frac{p}{2}} + \varepsilon^2\lambda} \tag{3.4.7}$$

for all $\lambda > 0$. The right-hand-side of (3.4.7) is minimized when $\lambda = \left(\frac{p}{2\varepsilon^2}\right)^{\frac{2}{2-p}}$ and by substituting the minimizer we get

$$\mathbb{P}\left(\left\{\omega \in \Omega : \|T_\omega(x)\|_{L_2} < \varepsilon \|x\|_{L_p}\right\}\right) \leq e^{-\frac{2}{\alpha}\left(\frac{p}{2}\right)^{\frac{\alpha}{p}} \frac{1}{\varepsilon^\alpha}} \tag{3.4.8}$$

where

$$\frac{1}{\alpha} = \frac{1}{p} - \frac{1}{2}. \tag{3.4.9}$$

If we let

$$C_p = \frac{2}{\alpha}\left(\frac{p}{2}\right)^{\frac{\alpha}{p}},$$

inequality (3.4.8) becomes

$$\mathbb{P}\left(\left\{\omega \in \Omega : \|T_\omega(x)\|_{L_2} < \varepsilon \|x\|_{L_p}\right\}\right) \leq e^{-\frac{C_p}{\varepsilon^\alpha}}$$

Were C_p is a constant depending only on p.

LEMMA 3.4.1 (Marcus-Pisier random embedding [19]). *For any finite sequence of vectors $x_1, x_2, \ldots, x_n \in L_p$, there exists an event $E \subset \Omega$ with probability at least $\frac{1}{2}$ such that for all $\omega \in E$*

$$\|x_i - x_j\|_{L_p} \leq K_p(\log n)^{\frac{1}{p}-\frac{1}{2}} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2} \tag{3.4.10}$$

Proof. Let $\varepsilon > 0$. For each pair of distinct indexes $i, j \in \{1, \ldots, n\}$ define

$$F_{ij} = \{\omega \in \Omega : \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2} < \varepsilon \|x_1 - x_j\|_{L_p}\}$$

By equation 3.4.8 we have that for all i, j

$$\mathbb{P}(F_{ij}) \leq e^{-\frac{C_p}{\varepsilon^\alpha}} \tag{3.4.11}$$

where C_p is a universal constant depending only on p. Define

$$F = \bigcup_{1 \leq i < j \leq n} F_{ij}.$$

Hence,

$$\mathbb{P}(F) \leq \binom{n}{2} e^{-\frac{C_p}{\varepsilon^\alpha}} \leq n^2 e^{-\frac{C_p}{\varepsilon^\alpha}}$$

Therefore we need to pick $\varepsilon > 0$ such that

$$n^2 e^{-\frac{C_p}{\varepsilon^\alpha}} < \frac{1}{2}$$

Hence, we can pick

$$\varepsilon = \left(\frac{C_p}{3 \log n}\right)^{\frac{1}{\alpha}}$$

for all $n \geq 2$. Substituting ε we have

$$\mathbb{P}(F) < \frac{1}{2}. \tag{3.4.12}$$

Hence, if $E = \Omega \setminus F$ then $\mathbb{P}(E) \geq \frac{1}{2}$ and whenever $\omega \in E$,

$$\|x_i - x_j\|_{L_p} \leq \left(\frac{3}{C_p}\right)^{\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}.$$

Using (3.4.9) we obtain (3.4.10). $\square$

We will know prove Theorem 3.3.3 using Lemma 3.4.1. We also use the fact that p-stable random variables have finite q absolute moment when $q < p$ (see equation 3.4.5). Unfortunately, this is the reason why this proof does not work for Conjecture 3.3.2.

Proof of Theorem 3.3.3. Suppose that $p \in (1, 2)$ and let $x_1, \ldots, x_n$ be a finite sequence of n vectors in L_p and A an $n \times n$ symmetric stochastic matrix. Let $t \in \mathbb{N}$ and $q < p$. With the same notation as in Theorem 3.4.1, there exists an event $E \subset \Omega$ with probability at least $1/2$ such that for all $\omega \in E$

$$\|x_i - x_j\|_{L_p} \leq K_p (\log n)^{\frac{1}{p} - \frac{1}{2}} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}.$$

Hence, for all $\omega \in E$

$$\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^q \leq K_p^q (\log n)^{\frac{q}{p}-\frac{q}{2}} \sum_{ij} A_{ij}^t \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q \qquad (3.4.13)$$

By Theorem 4.4 in [23], L_2 has Markov Type q for all $q < 2$. Hence, there is a constant C depending only on q such that

$$\sum_{ij} A_{ij}^t \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q \leq C^q t \sum_{ij} a_{ij} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q . \qquad (3.4.14)$$

Thus, for all $\omega \in E$,

$$\begin{aligned}
\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^q &\leq K_p^q (\log n)^{\frac{q}{p}-\frac{q}{2}} \sum_{ij} A_{ij}^t \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q \\
&\leq K_p^q (\log n)^{\frac{q}{p}-\frac{q}{2}} \sum_{ij} A_{ij}^t \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q \\
&\leq C^q K_p^q (\log n)^{\frac{q}{p}-\frac{q}{2}} t \sum_{ij} a_{ij} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q .
\end{aligned}$$

Integrating the last inequality over E we get

$$\begin{aligned}
\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^q &\leq C^q K_p^q (\log n)^{\frac{q}{p}-\frac{q}{2}} t \sum_{ij} a_{ij} \frac{1}{\mathbb{P}(E)} \int_E \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q \, d\mathbb{P} \\
&\leq 2C^q K_p^q (\log n)^{\frac{q}{p}-\frac{q}{2}} t \sum_{ij} a_{ij} \mathbb{E} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q
\end{aligned}$$

where the last inequality above is obtained since $P(E) \geq 1/2$.
On the other hand, by equality (3.4.5),

$$\delta_{qp}^q \|x\|_p^q = \mathbb{E}|T(x)|^q = \mathbb{E}_{\mathbb{P}} \mathbb{E}_{\mathbb{P}'} |T_{(\omega,\omega')}(x)|^q = \delta_{q2}^q \mathbb{E}_{\mathbb{P}}[\|T_\omega(x)\|^q].$$

for all $x \in L_p$. Hence, for some constant c_{pq} depending only on p and q,

$$\mathbb{E} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2}^q = c_{pq} \|x_i - x_j\|_{L_p}^q . \qquad (3.4.15)$$

Finally, by (3.4.15)

$$\sum_{ij} A_{ij}^t \|x_i - x_j\|_{L_p}^q \leq 2C^q K_p^q c_{pq} (\log n)^{\frac{q}{p} - \frac{q}{2}} t \sum_{ij} a_{ij} \|x_i - x_j\|_{L_p}^q .$$

□

3.4.1 The random walk on the hypercube

In the case when the matrix A is the transition matrix of the standard random walk on a hypercube we have the following result.

PROPOSITION 3.4.2. *Conjecture 3.3.2 is true for the standard random walk on the hypercube of any dimension.*

The proof of this proposition will use the following theorem in [23].

THEOREM 3.4.3. *[23] For every $p \in (1, \infty)$ and $q \in (1, 2]$ there is a constant $C(p, q)$ with the following properties. Let X be a normed space with modulus of smoothness of power of Type p. Then for every n-points $x_1, \dots, x_n \in X$, $n \times n$ symmetric stochastic matrix A and $t \in \mathbb{N}$, if $p \leq q$ then*

$$\sum A_{ij}^t \|x_i - x_j\|^p \leq C(p, q) S_q(X)^q t \sum a_{ij} \|x_i - x_j\|^p .$$

i.e., X has Markov Type p. If $p > q$, then

$$\sum A_{ij}^t \|x_i - x_j\|^p \leq C(p, q) S_q(X)^p t^{\frac{p}{q}} \sum a_{ij} \|x_i - x_j\|^p .$$

The proof of Proposition is simple given Theorem 3.4.3 and Theorem 6 in [25]. However, we included it here because we believe the standard random walk on the hypercube is sort of an extremal case of Conjecture 3.3.2.

Proof of Proposition 3.4.1. Let us denote P the transition probability matrix of the standard random walk on the Hypercube $\{-1, 1\}^n$. Let us also denote by P_t the

tth power of this matrix. One can see that $P_t(\varepsilon,\delta)$ only depends on the Hamming distance

$$\mathcal{H}(\varepsilon,\delta)=|\{i:\varepsilon_i\neq\delta_i\}|$$

so there are $n+1$ different values for P_t depending on the distances $\{0,1,....,n\}$. Let us denote them by $P_t(k)$ where $k\in 0,1,\ldots,n$. Hence, if $\varepsilon,\delta\in\{-1,1\}^n$ are such that

$$\mathcal{H}(\varepsilon,\delta)=k,$$

then $P_t(\varepsilon,\delta)=P_t(k)$. Hence,

$$\sum_{\delta,\varepsilon\in\{-1,1\}^n} P_t(\delta,\varepsilon)\,\|f(\delta)-f(\varepsilon)\|_{L_p}^2=\sum_{k=1}^{n}P_t(k)\sum_{\mathcal{H}(\delta,\varepsilon)=k}\|f(\delta)-f(\varepsilon)\|_{L_p}^2$$

On the other hand, if we denote by $\{C_i^k : k\in\{1,\ldots,2^{n-k}\binom{n}{k}\}$ the $2^{n-k}\binom{n}{k}$ k-dimensional faces of $\{-1,1\}^n$, then we have that

$$\sum_{\mathcal{H}(\delta,\varepsilon)=k}\|f(\delta)-f(\varepsilon)\|_{L_p}^2=\sum_{i=1}^{2^{n-k}\binom{n}{k}}\sum_{(\delta,\varepsilon)\in diag(C_i^k)}\|f(\delta)-f(\varepsilon)\|_{L_p}^2$$

Following the second remark after Theorem 6 in [25] we have that

$$\begin{aligned}\sum_{(\delta,\varepsilon)\in\text{diag}(C_i^k)}\|f(\delta)-f(\varepsilon)\|_{L_p}^2 &\precsim_p \sum_{\delta\in C_i^k}\left(\sum_{(\delta,\varepsilon)\in\text{edge}(C_i^k)}\|f(\delta)-f(\varepsilon)\|_{L_p}^p\right)^{\frac{2}{p}}\\ &\precsim_p k^{\frac{2}{p}-1}\sum_{(\delta,\varepsilon)\in\text{edge}(C_i^k)}\|f(\delta)-f(\varepsilon)\|_{L_p}^2\end{aligned}$$

Hence,

$$\sum_{(\delta,\varepsilon)\in\text{diag}(C_i^k)}\|f(\delta)-f(\varepsilon)\|_{L_p}^2\precsim_p k^{\frac{2}{p}-1}\sum_{(\delta,\varepsilon)\in\text{edge}(C_i^k)}\|f(\delta)-f(\varepsilon)\|_{L_p}^2$$

If we denote by C de number of times that we count twice every edge on $\{-1,1\}^n$

when we count them by counting the number of edges on each k-dimensional face, then we have by symmetry of the cube that

$$2^{k-1}\binom{k}{1}2^{n-k}\binom{n}{k} = C2^{n-1}\binom{n}{1}$$

and hence $C = \frac{k}{n}\binom{n}{k}$. Therefore we have that

$$\sum_{i=1}^{2^{n-k}\binom{n}{k}} \sum_{(\delta,\varepsilon)\in \text{edge}(C_i^k)} \|f(\delta)-f(\varepsilon)\|_{L_p}^2 = \frac{k}{n}\binom{n}{k} \sum_{\mathcal{H}(\delta,\varepsilon)=1} \|f(\delta)-f(\varepsilon)\|_{L_p}^2 .$$

Thus,

$$\begin{aligned}\sum_{\delta,\varepsilon\in\{-1,1\}^n} P_t(\delta,\varepsilon)\,\|f(\delta)-f(\varepsilon)\|_{L_p}^2 &\lesssim_p \sum_{k=1}^{n} P_t(k)\binom{n}{k}\frac{k^{\frac{2}{p}}}{n} \sum_{\mathcal{H}(\delta,\varepsilon)=1} \|f(\delta)-f(\varepsilon)\|_{L_p}^2 \\ &\lesssim_p n^{\frac{2}{p}} \sum_{\mathcal{H}(\delta,\varepsilon)=1} \frac{1}{n}\,\|f(\delta)-f(\varepsilon)\|_{L_p}^2\end{aligned}$$

If $t \leq n$ we can apply Theorem 3.4.3 and we obtain the desired result. $\square$

Remark 3.4.4. Proposition 3.4.1 remains true if one replaces the transition matrix of a standard random walk on the hypercube by any transition matrix on the hypercube for which the transition probabilities only depend on the Hamming distance between the vertices of the cube.

3.5 General Extension Theorems

In this section we state some general criteria to extend Lipschitz maps. The three extension theorems below are due to Ball [3]. The first theorem states that in order to extend a Lipschitz function defined on a metric space into a reflexive vector space, it is enough to know if we can extend it when restricted to finite subsets of the domain.

THEOREM 3.5.1. *[3] Let (X, d) be a metric space and Y a normed space, Z a subset of X and $f : Z \to Y$ Lipschitz. Then, there is an extension $F : X \to Y^{**}$ of f with $\|F\|_{Lip} \le K$ if and only if for every finite subset $S = \{x_1, \ldots, x_n\} \subset X$ there is a function F_S such that $F_S|_{S\cap Z} = f_{S\cap Z}$ and $\|F_S\|_{Lip} \le K$*

The second theorem gives us a criterion to extend Lipschitz functions defined on finitely many points in terms of positive symmetric matrices.

THEOREM 3.5.2. *[3] Let (X, d) be a metric space and Y a normed space, Z a subset of X and $f : Z \to Y$ Lipschitz. Given finite subset $S = \{x_1, \ldots, x_n\} \subset X$ there is a function F_S such that $F_S|_{S\cap Z} = f_{S\cap Z}$ and $\|F_S\|_{Lip} \le K$ if and only if for any $n \times n$ symmetric matrix $H = (h_{ij})$ with positive entries, there exists map $F^H : \{x_1, \ldots, x_n\} \to Y^{**}$ which agrees with f on $Z \cap \{x_1, \ldots, x_n\}$ and satisfies*

$$\sum_{ij} h_{ij} \left\|F^H(x_i) - F^H(x_j)\right\|^2 \le K^2 \sum_{ij} h_{ij} d(x_i, x_j)^2$$

Finally, the third theorem restated Theorem 3.5.2 in terms of symmetric stochastic matrices.

THEOREM 3.5.3. *[3] Let (X, d) be a metric space and Y a normed space, Z a subset of X and $f : Z \to Y$ Lipschitz. Given finite subset $S = \{x_1, \ldots, x_n\} \cup \{z_1, \ldots z_m\} \subset X$ where $z_1, \ldots, z_m \in Z$, there is a function F_S such that $F_S|_{S\cap Z} = f_{S\cap Z}$ and $\|F_S\|_{Lip} \le K$ if and only if for all $m, n \in \mathbb{N}$, $n \times n$ symmetric stochastic matrix A, $n \times m$ stochastic matrix B, and $\alpha \in (0, 1)$, there are points $(y_i)_{i=1}^n$ in Y satisfying*

$$\begin{aligned}\alpha \sum a_{ij} \|y_i - y_j\|^2 + 2(1-\alpha) \sum b_{ir} \|y_i - f(z_r)\|^2 \\ \le K^2 \left(\alpha \sum a_{ij} d(x_i, x_j)^2 + 2(1-\alpha) \sum b_{ir} d(x_i, z_r)^2\right)\end{aligned}$$

3.6 The Main Extension Theorems

DEFINITION 3.6.1. *The n-Markov Type p constant of a metric space $(\mathcal{M}, d_{\mathcal{M}})$, denoted by $M_p^n(\mathcal{M})$, is the smallest constant M such that for any $n, t \in \mathbb{N}$, any $n \times n$ symmetric stochastic matrix $A = (a_{ij})$, and $x_1, ..., x_n \in \mathcal{M}$,*

$$(1-\alpha)\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} d_{\mathcal{M}}(x_i, x_j)^p \leq \alpha M^p \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} d_{\mathcal{M}}(x_i, x_j)^p.$$

where $C = (1-\alpha)(I - \alpha A)^{-1}$. With this notation

$$M_p(\mathcal{M}) = \sup_n M_p^n(\mathcal{M})$$

Following the steps of the proof of Ball's Extension Theorem it can be easily shown the following.

THEOREM 3.6.2. *Let $\mathcal{M}$ be a a metric space and Y a vector space, then*

$$e^n(\mathcal{M}, Y) \precsim M_2^n(\mathcal{M}) N_2(Y)$$

Note that by Theorem 3.3.1, $M_2^n(\mathcal{M}) \precsim \sqrt{\log n}$ for any metric space $\mathcal{M}$. Thus, we have the following simple corollary.

COROLLARY 3.6.3. *Let $\mathcal{M}$ be a a metric space and Y a vector space, then*

$$e^n(\mathcal{M}, Y) \precsim \sqrt{\log n} N_2(Y).$$

Conjecture 3.2.11 can be restated as showing that

$$M_2^n(L_p) \precsim_p (\log n)^{\frac{1}{p}-\frac{1}{2}}.$$

This would imply, by a simple application of Theorem 3.6.2, that $e^n(L_p, Y) \precsim_p$

$(\log n)^{\frac{1}{p}-\frac{1}{2}} N_2(Y)$ for any vector space Y. We give a proof of this using Lemma 3.4.1.

THEOREM 3.6.4. *Let* $p \in (1,2)$. *Then*

$$e^n(L_p, Y) \precsim_p (\log n)^{\frac{1}{p}-\frac{1}{2}} N_2(Y)$$

The proof consists on putting to work together the machinery in [19] and [3] as we shall see below.

Proof. Let $X = \{x_1, \ldots, x_n\}$ be an n-point subset of L_p. Then, by Lemma 3.4.1 there exists a constant K_p depending on p and a subset E of Ω with probability $\mathbb{P}(E) \geq \frac{1}{2}$ such that for all $\omega \in E$ we have that

$$\|x_i - x_j\|_{L_p} \leq K_p (\log n)^{\frac{1}{p}-\frac{1}{2}} \|T_\omega(x_i) - T_\omega(x_j)\|_{L_2} \tag{3.6.1}$$

for all $i, j \in \{1, \ldots, n\}$.

Let $S = \{z_1, \ldots, z_m\} \cup \{x_1, \ldots, x_n\} \subset Z \cup X$ be an arbitrary finite subset such that $\{z_1, \ldots, z_m\} \subset Z$, A an $n \times n$ symmetric stochastic matrix, B and $n \times m$ stochastic matrix, and $\alpha \in (0,1)$. Let

$$y_i = \sum_r (CB)_{ir} f(z_r)$$

for $i \in \{1, \ldots, n\}$. Then, by the definition of $N_2(Y)$,

$$\begin{aligned}
&\alpha \sum a_{ij} \|y_i - y_j\|_2^2 + 2(1-\alpha) \sum b_{ir} \|y_i - f(z_r)\|_2^2 \\
&\leq (1-\alpha) N_2(Y) \sum c_{ij} \left\| \sum_r b_{ir} f(z_r) - \sum_s b_{js} f(z_s) \right\|_2^2 + 2(1-\alpha) \sum b_{ir} \|y_i - f(z_s)\|_2^2
\end{aligned}$$

By convexity of the function $x \to \|ax+b\|^2$, this is at most

$$3(1-\alpha)N_2(Y)\sum(B^TCB)_{rs}\left\|f(z_r)-f(z_s)\right\|_2^2. \tag{3.6.2}$$

Applying the Lipschitz condition implies that this is at most

$$3N_2(Y)\left\|f\right\|_{\text{Lip}}^2(1-\alpha)\sum(B^TCB)_{rs}\left\|z_r-z_s\right\|_p^2 \tag{3.6.3}$$

On the other hand we have that (3.6.3) is at most

$$\begin{aligned}
&(1-\alpha)N_2(Y)\sum(B^TCB)_{rs}\left\|z_r-z_s\right\|_p^2\\
&\qquad\le(1-\alpha)\sum_{ijrs}c_{ij}b_{ir}b_{js}(\|z_r-x_i\|+\|x_i-x_j\|+\|x_j-z_s\|)^2\\
&\le 3N_2(Y)K_p(\log n)^{\frac{2}{p}-1}(1-\alpha)\left[\sum c_{ij}\left\|T_\omega(x_i)-T_\omega(x_j)\right\|_{L_2}^2+2\sum b_{ir}\left\|x_i-z_r\right\|_p^2\right]\\
&\le 3N_2(Y)K_p(\log n)^{\frac{2}{p}-1}\left[\alpha\sum a_{ij}\left\|T_\omega(x_i)-T_\omega(x_j)\right\|_{L_2}^2+2(1-\alpha)\sum b_{ir}\left\|x_i-z_r\right\|_p^2\right]
\end{aligned}$$

for each $\omega\in E$, where the leftmost inequality above was obtain by using the fact that $\mathcal{M}_2(L_2)=1$. Let us define a family of new metrics on L_p by

$$d_\omega(x,y)=\sup\{\left\|T_\omega(x)-T_\omega(y)\right\|_{L_2},\left\|x-y\right\|_p\}$$

Hence this define a random variable $d(x,y)$ for each $x,y\in L_p$. Therefore, we have that for all $\omega\in\Omega_0$

$$\begin{aligned}
&(1-\alpha)\sum(B^TCB)_{rs}\left\|z_r-z_s\right\|_p^2\\
&\qquad\le 3N_2(Y)K_p(\log n)^{\frac{2}{p}-1}\left[\alpha\sum a_{ij}d_\omega(x_i,x_j)^2+2(1-\alpha)\sum b_{ir}d_\omega(x_i,z_r)^2\right]
\end{aligned}$$

By Theorem 3.5.3 for each $\omega\in E$ there exists a function $F_\omega^S: S\to L_2$ such that

$F_\omega^S|_{S\cap Z} = f|_{S\cap Z}$ and

$$\left\|F_\omega^S\right\|_{\text{Lip}} \leq 9K_p(\log n)^{\frac{1}{p}-\frac{1}{2}} \left\|f\right\|_{\text{Lip}}$$

where $\left\|F_\omega^S\right\|_{\text{Lip}}$ is the Lipschitz norm of F_ω^S with respect to the metric d_ω. Define $F^S : S \to L_2$ by

$$F^S(x) = \frac{1}{\mathbb{P}(E)} \int_E F_\omega^S(x) d\mathbb{P}(\omega) \tag{3.6.4}$$

for all $x \in S$. Hence, if $x, y \in S$, we have that

$$\begin{aligned}
\left\|F^S(x) - F^S(y)\right\|_2 &\leq \frac{1}{\mathbb{P}(E)} \int_E \left\|F_\omega^S(x) - F_\omega^S(y)\right\| d\mathbb{P}(\omega) \\
&\leq 9N_2(Y)K_p(\log n)^{\frac{1}{p}-\frac{1}{2}} \frac{\left\|f\right\|_{\text{Lip}}}{\mathbb{P}(E)} \int_E d_\omega(x,y) d\mathbb{P}(\omega) \\
&\leq 9N_2(Y)K_p(\log n)^{\frac{1}{p}-\frac{1}{2}} \frac{\left\|f\right\|_{\text{Lip}}}{2} \mathbb{E}_{\mathbb{P}}\left[d_\omega(x,y)\right].
\end{aligned}$$

On the other hand, we have that

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[d_\omega(x,y)\right] &\leq \mathbb{E}_{\mathbb{P}}\left[\left\|T_\omega(x) - T_\omega(y)\right\|_2 + \left\|x-y\right\|_p\right] \\
&= \mathbb{E}_{\mathbb{P}}\left[\left\|T_\omega(x) - T_\omega(y)\right\|_2\right] + \left\|x-y\right\|_p \\
&= (\delta_{1p}\delta_{12} + 1)\left\|x-y\right\|_p
\end{aligned}$$

and therefore

$$\left\|F^S(x) - F^S(y)\right\|_2 \precsim_p N_2(Y)(\log n)^{\frac{1}{p}-\frac{1}{2}} \left\|f\right\|_{\text{Lip}} \left\|x-y\right\|_p$$

The result follows by applying Theorem3.5.1. $\square$

3.7 Further Remarks

After this work was done, Mendel and Naor [21] informed us that they showed the following simple relation between finitary Lipschitz extension moduli

$$e^n(\mathcal{M}, \mathcal{N}) \leq 2 + e_n(\mathcal{M}, \mathcal{N}).$$

Hence, the results about the extension of Lipschitz maps in this chapter can be also obtain by using this relation and the very well developed theory for the constant $e_n(\mathcal{M}, \mathcal{N})$. As it was pointed out in [21] this inequality cannot be reversed in general. In fact, $e^n(\mathcal{M}, \mathcal{N}) \leq n + 1$ (see [8], theorem 1.1), but if $\mathcal{M} = \mathbb{R}$ and $\mathcal{N} = \{0, 1\}$ then $e_n(\mathcal{M}, \mathcal{N}) = \infty$, since all continuous functions from $\mathbb{R}$ to $\{0, 1\}$ are constant. It would be interesting to obtain a reverse inequality for particular sub-classes of metric spaces. For example, when $\mathcal{N} = Y$ for Y a Banach space. In such a case, our approach will imply several important bounds for the constant $e^n(\mathcal{M}, Y)$.

Bibliography

[1] G. Ambrus. *Analytic and probabilistic problems in discrete geometry*. PhD thesis, UCL (University College London), 2009.

[2] K. M. Ball. The plank problem for symmetric bodies. *Inventiones mathematicae*, 104(1):535–543, 1991.

[3] K. M. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geometric & Functional Analysis GAFA*, 2(2):137–172, 1992.

[4] K. M. Ball. Convex geometry and functional analysis. *Handbook of the geometry of Banach spaces*, 1:161–194, 2001.

[5] K. M. Ball. The complex plank problem. *Bulletin of the London Mathematical Society*, 33(4):433–442, 2001.

[6] K. M. Ball, O. Ortega-Moreno, and M. Prodromou. Hadamard matrices and 1-factorization of complete graphs. *Mathematika*, 65(3):488–499, 2019.

[7] T. Bang. A Solution of the "Plank Problem". *Proceedings of the American Mathematical Society*, 2(6):990–993, 1951.

[8] G. Basso. Lipschitz extensions to finitely many points. *Analysis and Geometry in Metric Spaces*, 6(1):174–191, 2018.

[9] A. Benedek and R. Panzone. The space L^p, with mixed norm. *Duke Mathematical Journal*, 28(3):301–324, 1961.

[10] K. Bezdek. Tarskis plank problem revisited. In *Geometryintuitive, discrete, and convex*, pages 45–64. Springer, 2013.

[11] P. Borwein and T. Erdélyi. *Polynomials and polynomial inequalities*, volume 161. Springer Science & Business Media, 1995.

[12] H. Davenport. *The Higher Arithmetic: An Introduction to the Theory of Numbers*. Cambridge University Press, 8 edition, 2008.

[13] Z. Jiang and A. Polyanskii. Proof of László Fejes Tóths zone conjecture. *Geometric and Functional Analysis*, 27(6):1367–1377, 2017.

[14] W. Johnson and J. Lindenstrauss. Extensions of Lipschhitz maps into a Hilbert space. *Contemporary Math*, 26, 1984.

[15] J. P. Kahane. Sur les sommes vectorielles $\sum \pm u_n$. *Comptes Rendus Hebdomadaires des Séances de l'Academie des Sciences*, 259(16):2577–2580, 1964.

[16] A. Khintchine. Über einen satz der wahrscheinlichkeitsrechnung. *Fundamenta Mathematicae*, 6(1):9–20, 1924.

[17] H. König, J. R. Retherford, and N. Tomczak-Jaegermann. On the eigenvalues of (p, 2)-summing operators and constants associated with normed spaces. *Journal of Functional Analysis*, 37(1):88–126, 1980.

[18] Y. J. Leung, W. V. Li, and Rakesh. The d-th linear polarization constant of $\mathbb{R}^d$. *Journal of Functional Analysis*, 255(10):2861–2871, 2008.

[19] M. Marcus and G. Pisier. Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. *Acta mathematica*, 152(1):245–301, 1984.

[20] M. Mendel and A. Naor. Spectral calculus and Lipschitz extension for barycentric metric spaces. *Analysis and Geometry in Metric Spaces*, 1:163–199, 2013.

[21] M. Mendel and A. Naor. A relation between finitary lipschitz extension moduli. *arXiv preprint arXiv:1707.07289*, 2017.

[22] P. Moree. *On primes in arithmetic progression having a prescribed primitive root*. Department of Mathematics, Computer Science, Physics and Astronomy, University of Amsterdam, 1999.

[23] A. Naor, Y. Peres, O. Schramm, and S. Sheffield. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Mathematical Journal*, 134(1):165–197, 2006.

[24] A. Naor and Y. Rabani. On Lipschitz extension from finite subsets. *Israel Journal of Mathematics*, 219(1):115–161, 2017.

[25] A. Naor and G. Schechtman. Remarks on non linear type and Pisier's inequality. *Journal fur die Reine und Angewandte Mathematik*, 552:213, 2002.

[26] F. Nazarov. The Bang solution of the coefficient problem. *Algebra i Analiz*, 9(2):272–287, 1997.

[27] R. Paley. On orthogonal matrices. *Journal of Mathematics and Physics*, 12(1-4):311–320, 1933.

[28] G. Pisier. Estimations des distances à un espace euclidien et des constantes de projection des espaces de Banach de dimension finie; d'après H. König et al. *Séminaire Analyse fonctionnelle(1978-1979)*, pages 1–21, 1979.

[29] L. F. Tóth. Exploring a planet. *The American Mathematical Monthly*, 80(9):1043–1044, 1973.